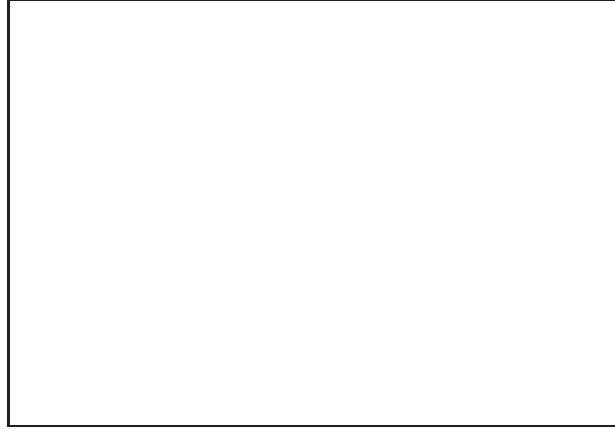


CORRELATIONS, BELL INEQUALITY VIOLATION & QUANTUM ENTANGLEMENT



A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
AT THE UNIVERSITY OF QUEENSLAND IN
JANUARY 2008

Yeong-Cherng Liang
Department of Physics

© Yeong-Cherng Liang, 2009.

Typeset in L^AT_EX 2_ε.

*To my beloved family members in Malaysia
and my dearest Shih-yin,
for their continuous support, encouragement and love ...*

“For those who are not shocked when they first come across quantum theory cannot possibly have understood it”, Niels Bohr, first quoted by Werner Heisenberg in Physics and Beyond, pp. 206 (New York: Harper & Row, 1971).

“When the ‘system’ in question is the whole world where is the ‘measurer’ to be found? Inside, rather than outside, presumably. What exactly qualifies some subsystems to play this role? Was the world wave function waiting to jump for thousands of millions of years until a single-celled living creature appeared? Or did it have to wait a little longer for some highly qualified measurer — with a Ph.D.?”, John S. Bell in Quantum Gravity 2, pp. 611 (1981).

“... we have always had a great deal of difficulty understanding the world view that quantum mechanics represents. At least I do, because I’m an old enough man that I haven’t got to the point that this stuff is obvious to me. Okay, I still get nervous with it ... you know how it always is, every new idea, it takes a generation or two until it becomes obvious that there’s no real problem. I cannot define the real problem, therefore I suspect there’s no real problem, but I’m not sure there’s no real problem”, Richard P. Feynman in International Journal of Theoretical Physics, **21**, pp. 467 (1982).

“We often discussed his notions on objective reality. I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it. The rest of this walk was devoted to a discussion of what a physicist should mean by the term ‘to exist’ ”, A. Pais, Reviews of Modern Physics, **51**, pp. 908 (1979).

“Bells theorem is the most profound discovery of science”, Henry P. Stapp, IL Nuovo Cimento, **29B**, pp. 271 (1975).

“Anybody who’s not bothered by Bell’s theorem has to have rocks in his head”, anonymous Princeton physicist, first quoted by N. David Mermin in Physics Today, **38**, pp. 41 (1985).

Statement of Originality

I hereby declare that, except where acknowledged below in the Statement of Contribution to Jointly-published Work and at other appropriate places in the thesis, the work presented in this thesis is original and my own work, and has not been submitted in whole or part for a degree in any university.

Statement of Contribution to Jointly-published Work

The content of Chapter 5 is based largely on Ref. [1], which results from a joint research project between me and my supervisor Andrew C. Doherty. I have done all the research in this chapter guided by initial ideas of Andrew C. Doherty and weekly discussion about the project. I wrote the initial draft of Ref. [1].

The content of Chapter 6 is based largely on Ref. [2], which results from a joint research project between me and my supervisor Andrew C. Doherty. Except for Sec. 6.2.2, which consists mostly of review material, I have done all the research in this chapter guided by initial ideas of Andrew C. Doherty and weekly discussion about the project. I wrote the initial draft of Ref. [2].

The content of Sec. 7.4 and Appendix A are based largely on Ref. [3] and Ref. [4], which result from a joint research project between me and my collaborators Lluís Masanes and Andrew C. Doherty. Ref. [3] and Ref. [4] were initially drafted, respectively, by Lluís Masanes and me. This project was initiated by Lluís Masanes who also did many of the initial calculations. The present form of Lemma 1 and Lemma 2 in Ref. [3] is a result of collaborations between the three of us. Results presented in Sec. II of Ref. [4] were obtained by Lluís Masanes and me, guided by weekly discussion with Andrew C. Doherty. Results presented in Sec. III C of Ref. [4] was initially obtained by Lluís Masanes and independently verified by me. Results presented in Sec. IV of Ref. [4] were initially obtained by me and independently verified by Lluís Masanes.

Yeong-Cherng Liang
Author

Dr. Andrew C. Doherty
Principal Advisor

Acknowledgements

This thesis is a consequence of direct and indirect contributions from various people, without whom I would not have gone this far. It is almost inevitable that I would miss some names in the following enumeration. For that matter, I would now declare that if you are in doubt, then yes, you must have been one of the contributors and I thank you for your help in one way or another. This is, of course, not an excuse for me to not express my appreciation explicitly and I shall attempt to do that in what follows.

The successful completion of my PhD candidature as well as this thesis would have been impossible, if not highly improbable without the help and guidance from my principal supervisor, Dr. Andrew C. Doherty. Andrew is always full of ideas and this means a lot to a junior researcher like me who sometimes lacks insight into the key issue of a problem. In particular, his physical intuition has enabled me to see the *forest*, instead of *trees* at various occasions. His comments on language usage have always been of great help too. On the other hand, I must also thank Andrew for his generous support in regard of me traveling overseas to attend conferences — these opportunities have, no doubt, greatly expanded my horizons. Finally, I owe Andrew a big thank you for going the extra miles to read through the earlier drafts of this thesis and giving me his valuable comments.

Many thanks to Guifré Vidal who has effectively acted as my associate supervisor, giving me his timely advice both in and out of Physics. This proved to be of utmost importance especially towards the end of my PhD candidature. Thanks also to Michael Nielsen, who has offered some critical comments on my research and who has always tried to made the quantum information science initiative in the University of Queensland a wonderful learning environment. His insistence on our active participation in seminars has undeniably done me a great favor in these years.

I am also thankful to my other present collaborators — Lluís Masanes, Ben Toner, Stephanie Wehner, Valerio Scarani — as well as past collaborators — Dagomir Kaszlikowski, Leong Chuan Kwek, Berthold-Georg Englert, Ajay Gopinathan and Choo Hiap Oh — for giving me the opportunity to learn from them. Among which, I am above all grateful to Dagomir Kaszlikowski and Leong Chuan Kwek for bringing me into this exciting field of quantum information science. Thanks also to Jing-Ling Chen, Shiang Yong Looi and Meng Khoon Tey for their helpful discussions.

The fellow PhD students sharing the same office as me are certainly not to be forgotten. I thank Eric Cavalcanti for enriching my philosophical understanding of nature, making this doctorate of philosophy a well-justified one. Of course, his never ending list of puzzles and paradoxes has also refreshed my ordinary research life from time to time. Chris Foster is definitely the most reliable mathematician, technician, and entertainer in Room 302. His

trademark of trying to make sloppy calculations in Physics rigorous is both enlightening and inspiring. His help over these years, which often results in delaying his own research progress, is greatly appreciated. Next, I would like to thank Paulo Mendonça for always willing to listen to my complaints and sharing with me the unusual experience that he has been through. Thanks also to Paulo for proofreading an earlier draft of this thesis. Certainly, the random humor from Andy Ferris is much appreciated too.

I am also grateful to all of my other friends, especially those coming from the UQ Badminton Club. You guys have complemented my academic life in just the right way. Finally, I would also like to acknowledge the generous financial support from the International Postgraduate Research Scholarships (IPRS) and the University of Queensland Graduate School Scholarships (UQGSS).

List of Publications

Publications by the Candidate Relevant to the Thesis

- Yeong-Cherng Liang and Andrew C. Doherty, *Better Bell-inequality violation by collective measurements*. Physical Review A **73**, 052116 (2006)
- Yeong-Cherng Liang and Andrew C. Doherty, *Bounds on quantum correlations in Bell-inequality experiments*. Physical Review A **75**, 042103 (2007)
- Lluís Masanes, Yeong-Cherng Liang and Andrew C. Doherty, *All bipartite entangled states display some hidden nonlocality*. Physical Review Letters **100**, 090403 (2008)
- Yeong-Cherng Liang, Lluís Masanes and Andrew C. Doherty, *Convertibility between two-qubit states using stochastic local quantum operations assisted by classical communication*. Physical Review A **77**, 012332 (2008)

Additional Publications by the Candidate Relevant to the Thesis but not Forming Part of it

- Andrew C. Doherty, Yeong-Cherng Liang, Stephanie Wehner, and Ben Toner, *The quantum moment problem and bounds on entangled multi-prover games*, Proceedings of the 23rd IEEE Conference on Computational Complexity, pp. 199–210 (eprint arXiv:0803.4373)

Abstract

It is one of the most remarkable features of quantum physics that measurements on spatially separated systems cannot always be described by a *locally causal theory*. In such a theory, the outcomes of local measurements are determined in advance solely by some unknown (or hidden) variables and the choice of local measurements. Correlations that are allowed within the framework of a locally causal theory are termed *classical*. Typically, the fact that quantum mechanics does not always result in classical correlations is revealed by the violation of *Bell inequalities*, which are constraints that have to be satisfied by any classical correlations. It has been known for a long time that *entanglement* is necessary to demonstrate nonclassical correlations, and hence a Bell inequality violation. However, since some entangled quantum states are known to admit explicit locally causal models, the exact role of entanglement in Bell inequality violation has remained obscure. This thesis provides both a comprehensive review on these issues as well as a report on new discoveries made to clarify the relationship between entanglement and Bell inequality violation. In particular, within the framework of a standard Bell experiment, i.e., a Bell inequality test that is directly performed on a single copy of a quantum state ρ , we have derived two algorithms to determine, respectively, a lower bound and an upper bound on the strength of correlations that ρ can offer for any given Bell inequality. Both of these algorithms make use convex optimization techniques in the form of a *semidefinite program*. By examples, we show that these algorithms can often be used in tandem, in conjunction with *convexity* arguments, to determine if a quantum state can offer nonclassical correlations and hence violates a given Bell inequality. On the other hand, since a standard Bell experiment typically involves measurements over many copies of the quantum systems, we have also investigated the possibility of enhancing the strength of nonclassical correlation by, instead, performing collective measurements on multiple copies of the quantum systems. Our findings show that even without postselection, such joint measurements may also lead to stronger nonclassical correlations, and hence a better Bell inequality violation. Meanwhile, previous studies have indicated that entangled state admitting locally causal models may still lead to observable nonclassical correlations if, prior to a standard Bell experiment, the state is subjected to some appropriate local preprocessing. This phenomenon of *hidden nonlocality* was discovered more than a decade ago, but to date, it is still not known if all entangled states can demonstrate nonclassical correlations through these more sophisticated Bell experiments. A key result in this thesis then consists of showing that for all bipartite entangled states, observable nonclassical correlations, in the form of a Bell-CHSH inequality violation, can indeed be derived if we allow both local preprocessing and the usage of shared ancillary state which by itself does not violate the Bell-CHSH inequality. This establishes a kind of equivalence between bipartite

entanglement and states that cannot be simulated by classical correlations. In summary, for a standard Bell experiment where no local preprocessing on a quantum state ρ is allowed, we have provided two algorithms that can be used in tandem to determine if ρ can be simulated by a locally causal theory, whereas in the scenario where local preprocessing is allowed, we have demonstrated that bipartite entangled states are precisely those which cannot always be simulated classically.

Contents

Acknowledgements	iii
List of Publications	v
Abstract	vii
List of Figures	xiii
List of Tables	xvii
List of Abbreviations	xix
1 Introduction	1
2 Bell's Theorem and Tests of Local Causality	5
2.1 Bell's Theorem	5
2.1.1 The Einstein-Podolsky-Rosen Incompleteness Arguments	5
2.1.2 Completeness and Hidden-Variable Theory	7
2.1.3 Quantum Mechanics is not a Locally Causal Theory	8
2.2 Towards an Experimental Test of Local Causality	11
2.2.1 Bell-Clauser-Horne-Shimony-Holt Inequality	11
2.2.2 Bell-Clauser-Horne Inequality	13
2.2.3 Experimental Progress	16
3 Classical Correlations and Bell Inequalities	17
3.1 Classical Correlations and Probabilities	17
3.2 Geometrical Structure of the Set of Classical Correlations	20
3.2.1 The Spaces of Correlations	20
3.2.2 The Convex Set of Classical Correlations	21
3.2.3 Correlation Polytope and Bell Inequalities	22
3.3 The Zoo of Bell Inequalities	26
3.3.1 Other Bipartite Bell Inequalities for Probabilities	26
3.3.1.1 Two Outcomes $\mathbf{n} = (2, 2)$	26
3.3.1.2 More than Two Outcomes	29
3.3.2 Other Bipartite Correlation Inequalities	30
3.3.3 Multipartite Bell Inequalities	32

3.4	Conclusion	34
4	Quantum Correlations and Locally Causal Quantum States	35
4.1	Introduction	35
4.2	Quantum Correlations	36
4.2.1	General Structure of the Set of Quantum Correlations	37
4.2.2	Quantum Correlation and Bell Inequality Violation	38
4.3	Locally Causal Quantum States	39
4.3.1	Separable States	40
4.3.2	Quantum States Admitting General LHVM	40
4.3.2.1	$U \otimes U$ Invariant States — Werner States	41
4.3.2.2	$U \otimes \overline{U}$ Invariant States — Isotropic States	42
4.3.2.3	$U \otimes U \otimes U$ Invariant States	44
4.3.3	Quantum States Satisfying Some Bell Inequalities	45
4.3.3.1	PPT Entangled States	45
4.3.3.2	Entangled States with Symmetric Quasiextension	47
4.4	Conclusion	48
5	Bounds on Quantum Correlations in Standard Bell Experiments	49
5.1	Introduction	49
5.2	Bounds on Quantum Correlations	51
5.2.1	Preliminaries	51
5.2.2	Algorithm to Determine a Lower Bound on $\mathcal{S}_{\text{QM}}(\rho)$	53
5.2.2.1	General Settings	53
5.2.2.2	Iterative Semidefinite Programming Algorithm	54
5.2.2.3	Two-outcome Bell Experiment	55
5.2.3	Algorithm to Determine an Upper Bound on $\mathcal{S}_{\text{QM}}(\rho)$	56
5.2.3.1	Global Optimization Problem	57
5.2.3.2	State-independent Bound	59
5.2.3.3	State-dependent Bound	60
5.2.3.4	Higher Order Relaxations	61
5.3	Applications & Limitations of the Two Algorithms	62
5.3.1	Bell-CHSH violation for Two-Qudit States	62
5.3.2	Bell-CH violation for Two-Qubit States	63
5.3.3	I_{3322} -violation for a Class of Two-Qubit States	65
5.3.4	Limitations of the UB algorithm	67
5.4	Conclusion	68
6	Bell-Inequality Violations by Quantum States	69
6.1	Introduction	69
6.2	Single Copy Bell Inequality Violation	71
6.2.1	Bell-CH-violation for Pure Two-Qudit States	71
6.2.2	CGLMP and I_{22nn} -violation for Some Two-Qudit States	74
6.3	Better Bell-inequality Violation by Collective Measurements	76
6.3.1	Multiple Copies of Pure States	76

6.3.2	Multiple Copies of Mixed States	79
6.4	Conclusion	82
7	Nonstandard Bell Experiments and Hidden Nonlocality	85
7.1	Introduction	85
7.2	Single Copy Nonstandard Bell Experiments	86
7.2.1	Nonstandard Bell Experiments on Pure Entangled States	86
7.2.2	Nonstandard Bell Experiments on Mixed Entangled States	87
7.2.2.1	Nonlocality Hidden in Werner States	88
7.2.2.2	Nonlocality Hidden in Standard Bell-CHSH Experiment	89
7.2.3	Justification of Single Copy Nonstandard Bell Experiment	90
7.3	Nonstandard Bell Experiments on Multiple Copies	91
7.3.1	Nonstandard Bell Experiments with Collective Measurements	92
7.3.2	Nonstandard Bell Experiments and Distillability	92
7.4	Observable Nonlocality for All Bipartite Entangled States	93
7.4.1	Bipartite States with No Bell-CHSH Violation after SLOCC	93
7.4.2	Nonstandard Bell Experiment with Shared Ancillary State	96
7.5	Conclusion	99
8	Conclusion	101
	References	105
A	Bell-diagonal Preserving Separable Maps	121
A.1	Four-qubit Separable States with $U \otimes U \otimes V \otimes V$ Symmetry	121
A.2	Separable Maps and SLOCC	126
A.3	Bell-diagonal Preserving SLOCC Transformations	127
B	Some Miscellaneous Calculations	129
B.1	Classical Correlations and Bell's Theorems	129
B.1.1	Equivalence between the CGLMP and I_{22nn} inequality	129
B.2	Quantum Correlations and Locally Causal Quantum States	134
B.2.1	Convexity of Non-Bell-Inequality-Violating States	134
B.3	Bounds on Quantum Correlations in Standard Bell Experiments	135
B.3.1	Bell-CH Inequality and Full Rank Projector	135
B.3.2	Derivation of Horodecki's Criterion using LB	136
B.4	Bell-Inequality Violations by Quantum States	138
B.4.1	Bell-CH-violation for Pure Two-Qudit States	138
B.5	Nonstandard Bell Experiments and Hidden Nonlocality	139
B.5.1	Proof of Lemma 17	139
C	Semidefinite Programming and Relaxations	143
C.1	Semidefinite Programs	143
C.2	Semidefinite Relaxation to Finding $\mathcal{S}_{\text{QM}}(\rho)$	144
C.2.1	Lowest Order Relaxation with Observables of Fixed Trace	145

C.2.2	Sufficient Condition for No-violation of the Bell-CHSH Inequality . .	146
C.3	Explicit Forms of Semidefinite Programs	147
C.3.1	SDP for the LB Algorithm	147
C.3.2	SDP for the UB Algorithm	148
C.3.2.1	State-independent Bound	148
C.3.2.2	State-dependent Bound	149
C.3.3	SDP for the Verification of Entanglement Witness	150
List of Symbols		151

List of Figures

2.1	Schematic diagram of the experimental setup involved in a standard two-party Bell experiment. The source produces pairs of physical systems that are subsequently distributed, respectively, to Alice and Bob. They then subject the physical system that they receive to an analyzer which has an adjustable parameter (denoted by α and β correspondingly). For example, in the case of polarization measurement on photons, an analyzer is simply a combination of waveplates and a polarizer. The final stage of the measurement process consists of detecting the subsystems that pass through each analyzer with one or more detectors. In the scenario considered by Bell [5] and Clauser <i>et al.</i> [6], there are two detectors at each site, whereas in the original experimental scenario considered by CH [7], there is only one detector after each analyzer.	14
3.1	Schematic representation of the LHVM corresponding to a particular extreme point of $\mathcal{P}_{3;4}^{6;6}$, denoted by $\mathbf{a}, \mathbf{b} \mathbf{B}_{AB}$, where $\mathbf{a} \equiv (\vartheta_1^{[1]}, \vartheta_2^{[1]}, \vartheta_3^{[1]}) = (2, 5, 3)$ and $\mathbf{b} \equiv (\vartheta_1^{[2]}, \vartheta_2^{[2]}, \vartheta_3^{[2]}, \vartheta_4^{[2]}) = (4, 5, 3, 6)$ (adapted from Figure 1 of Ref. [8]). Each row (column), separated from each other by solid horizontal (vertical) lines, corresponds to a choice of measurement s_a (s_b) for Alice (Bob). The intersection of a row and a column gives rise to a sector, which corresponds to particular choice of Alice's and Bob's measurement. For each extremal LHVM, the outcome of measurements solely depends on the choice of local measurement. Hence, once a row (column) is chosen, the measurement outcome is also determined, and is indicated by a dashed horizontal (vertical) line. For example, Alice will always observe the second outcome ($o_a = 2$) whenever she chooses to perform the first measurement ($s_a = 1$), regardless of Bob's choice of measurement.	23

- 4.1 Plot of the various threshold weights p_{S,W_d} , p_{L,W_d}^Π , p_{L,W_d}^{POVM} for Werner states $\rho_{W_d}(p)$, and p_{S,I_d} , p_{L,I_d}^Π , p_{L,I_d}^{POVM} for isotropic states $\rho_{I_d}(p)$ as a function of d . Notice that for these two classes of states, the threshold weights for separability, i.e., p_{S,W_d} and p_{S,I_d} are identical; likewise for the threshold weights whereby a LHV for POVM measurements is known to exist, i.e., p_{L,W_d}^{POVM} and p_{L,I_d}^{POVM} . For each d , the vertical line joining these two threshold weights, which are, respectively, marked by a red + and a black \times , correspond to weights p of $\rho_{W_d}(p)$ and $\rho_{I_d}(p)$ whereby the states are entangled but do not violate any Bell inequalities. Similarly, for each d , the vertical line joining a blue \square and a red + corresponds to $\rho_{W_d}(p)$ which are entangled but are NBIV with projective measurements whereas the vertical line joining a purple circle and a red + corresponds to $\rho_{I_d}(p)$ that are entangled but are NBIV with projective measurements. 43
- 5.1 Domains of p where the compatibility between a locally causal description and quantum mechanical prediction given by $\rho_{CG}(p)$ was studied via the LB and UB algorithms in conjunction with the I_{3322} inequality. From right to left are respectively the domain of p whereby $\rho_{CG}(p)$ is: (D) found to violate the I_{3322} inequality; (C) found to give a lowest order upper bound that is compatible with the I_{3322} inequality; (B) found to give a higher order upper bound that is compatible with the I_{3322} inequality; (A) not known if it violates the I_{3322} inequality. 66
- 5.2 Numerical upper bound on $\mathcal{S}_{QM}^{(CH)}(\rho_H)$ obtained from the UB algorithm using lowest order relaxation and Eq. (5.45). The dashed horizontal line is the threshold above which no locally causal description is possible. 67
- 6.1 Best known Bell-CH inequality violation of pure two-qubit states obtained from Eq. (6.3), plotted as a function of ϕ , which gives a primitive measure of entanglement; $\phi = 0$ for bipartite pure product state and $\phi = 45^\circ$ for bipartite maximally entangled state. The curves from right to left represent increasing numbers of copies. The dotted horizontal line at $\frac{1}{\sqrt{2}} - \frac{1}{2}$ is the maximal possible violation of Bell-CH inequality; correlations allowed by locally causal theories have values less than or equal to zero. The solid line is the maximal Bell-CH inequality violation of $|\Phi_2\rangle$ determined using the Horodecki criterion, c.f. Appendix B.3.2. 77
- 6.2 Best known expectation value of the Bell operator coming from the Bell-CH inequality [\mathcal{B}_{CH} , Eq. (5.39)], I_{3322} inequality [$\mathcal{B}_{I_{3322}}$, Eq. (5.43)], and the I_{2244} inequality [$\mathcal{B}_{I_{2244}}$, Eq. (6.9)] with respect to the 2-dimensional Werner states $\rho_{W_2}(p)$; p represents the weight of the spin- $\frac{1}{2}$ singlet state in the mixture. Note that the best I_{2244} -violation found for $|\Psi^-\rangle^{\otimes 2}$ agrees with the best known violation presented in Table 6.1. Also included is the upper bound on $\mathcal{S}_{QM}^{(CH)}(\rho_{W_2}^{\otimes 2})$ obtained from the UB algorithm. 80

6.3	Distribution of two-qubit states sampled for better Bell-CH violation by collective measurements. The maximally entangled mixed states (MEMS), which demarcate the boundary of the set of density matrices on this concurrence-entropy plane [9, 10], are represented by the solid line. Note that as a result of the chosen distribution over mixed states this region is not well sampled. The region bounded by the solid line and the horizontal dashed line (with concurrence equal to $1/\sqrt{2}$) only contain two-qubit states that <i>violate</i> the Bell-CH inequality [11]; the region bounded by the solid line and the vertical dashed line (with normalized linear entropy equal to $2/3$) only contain states that <i>do not violate</i> the Bell-CH inequality [11, 12, 13]. Two-qubit states found to give better 3-copy Bell-CH violation are marked with red crosses.	81
6.4	Best known expectation value of the Bell operator coming from the Bell-CH inequality [\mathcal{B}_{CH} , Eq. (5.39)] and the I_{3322} inequality [$\mathcal{B}_{I_{3322}}$, Eq. (5.43)], with respect to the 3-dimensional isotropic states, $\rho_{I_3}(p)$; p is the weight of maximally entangled two-qutrit state in the mixture. The single copy Bell-CH inequality violation found here through LB is identical with the maximal violation, $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{I_3})$, found by Ito <i>et al.</i> in Ref. [14].	82
7.1	The relevant parameter space for $\rho_G(p, \theta)$. The set of states that do not violate the Bell-CHSH inequality but which do after the local filtering operations given by Eq. (7.10) is the shaded region bounded by the black dashed line ($p = p'_L$), the blue dotted line ($p = p_0$) and the red dotted line ($p = p_L$). . .	90
7.2	Schematic diagram illustrating the local filtering operations $\tilde{F}_{\mathcal{A}}$ and $\tilde{F}_{\mathcal{B}}$ involved in our protocol. The solid box on top is a schematic representation of the state σ whereas that on the bottom is for the ancilla state ρ . Left and right dashed boxes, respectively, enclose the subsystems possessed by the two experimenters \mathcal{A} and \mathcal{B}	97
A.1	A schematic diagram for the subsystems constituting ρ . Subsystems that are arranged in the same row in the diagram have $U \otimes U$ symmetry and hence are represented by Bell-diagonal states [15] (see text for details). In this Appendix, we are interested in states that are separable between subsystems enclosed in the two dashed boxes.	122

List of Tables

- 3.1 A hypothetical set of experimental data gathered in an experiment to test the Bell-CHSH inequality or the Bell-CH inequality. Here, n is an index to label each run of the experiment and N is some very large number such that the data set is statistically significant. The local measurements that may be performed by Alice are labeled by A_1 and A_2 whereas that for Bob are labeled by B_1 and B_2 . Outcomes of the experiments are labeled by ± 1 and are tabulated under the respective local measurements that are carried out in each run of the experiment. 19
- 3.2 The same set of experimental data as in Table 3.1 but with the *unperformed* measurement results (enclosed within round brackets) filled in according to some hypothetical LHVM. In particular, the LHVM works in such a way that the newly filled entries in the table give rise to the same joint and marginal probabilities as the original entries listed in Table. 3.1, c.f. Eq. (3.1). 19
- 5.1 The various threshold values for isotropic states $\rho_{I_d}(p)$. The first column of the table is the dimension of the local subsystem d . From the second column to the seventh column, we have, respectively, the value of p below which the state is separable p_{s,I_d} ; the value of p below which Eq. (5.35) is satisfied $p_{\text{UB-semianalytic}}$, and hence the state does not violate the Bell-CHSH inequality; the value of p below which the upper bound obtained from lowest order relaxation is compatible with Bell-CHSH inequality; the value of p below which the state cannot violate any Bell inequality via projective measurements; the value of p below which the state cannot violate any Bell inequality (Sec. 4.3.2.2); and the value of p above which a Bell-CHSH violation has been observed using the LB algorithm. 63

- 6.1 Best known CGLMP-violation and I_{22dd} -violation for the maximally entangled two-qudit state $|\Phi_d^+\rangle$. The first column of the table gives the dimension of the local subsystem d . The second column gives the largest possible quantum violation of the CGLMP inequality for $d \leq 8$, first obtained in Ref. [16], and subsequently verified in Ref. [17]; these maximal violations also set an upper bound on the maximal violation attainable by $|\Phi_d^+\rangle$ for each d . The third column of the table gives the best known d -outcome CGLMP-violation for $|\Phi_d^+\rangle$ whereas the fourth column gives the corresponding best known I_{22dd} -violation obtained from Eq. (6.8). Also included in the fifth column of the table is the threshold weight p_d below which no violation of either inequality by isotropic state $\rho_{I_d}(p)$ is known. 75
- 6.2 Best known Bell-CH inequality violation for some bipartite pure entangled states, obtained from Eq. (6.2) and Eq. (B.21) with and without collective measurements. Also included below is the upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(|\Phi\rangle\langle\Phi|)$ obtained from the UB algorithm. Each of these upper bounds is marked with a † . The first column of the table gives the number of copies N involved in the measurements. Each quantum state is labeled by its non-zero Schmidt coefficients, which are separated by : in the subscripts attached to the ket vectors; e.g., $|\Phi\rangle_{3:3:2:1}$ is the state with unnormalized Schmidt coefficients $\{c_i\}_{i=1}^4 = \{3, 3, 2, 1\}$. For each quantum state there is a box around the entry corresponding to the smallest N such that the lower bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(|\Phi\rangle\langle\Phi|)$ exceeds the single-copy upper bound (coming from the UB algorithm or otherwise). A violation that is known to be maximal is marked with a $*$ 78

List of Abbreviations

aka also known as

lhs left-hand-side

rhs right-hand-side

BIV Bell-inequality-violating

CH Clauser-Horne

CHSH Clauser-Horne-Shimony-Holt

CGLMP Collins-Gisin-Linden-Massar-Popescu

CPM completely positive map

EPR Einstein-Podolsky-Rosen

GHZ Greenberger-Horne-Zeilinger

LB lower bound

LHV local hidden variable

LHVM local hidden-variable model

LHVT local hidden-variable theory

LMI linear matrix inequality

LOCC local quantum operations assisted by classical communication

MEMS maximally entangled mixed states

NBIV non-Bell-inequality-violating

NSD negative semidefinite

POVM positive-operator-valued measure

PPT positive-partial-transposed

PSD	positive semidefinite
QCQP	quadratically-constrained quadratic program
SLO	stochastic local quantum operations without communication
SLOCC	stochastic local quantum operations assisted by classical communication
SDP	semidefinite program
SOS	sum of squares
UB	upper bound

1

Introduction

The advent of *Quantum Mechanics* is undeniably an important milestone in our attempt to understand Nature. On the one hand, quantum mechanics is well-known for giving very accurate predictions for microscopic phenomena, whereas on the other, it has also given some counter-intuitive predictions which seem nonsensical from a classical view point. Among the many intriguing features of quantum mechanics is *entanglement* [18, 19] which, loosely speaking, refers to the situation whereby two or more spatially separated physical systems are so strongly correlated that it may become impossible to independently describe the *physical state* of the individual systems. The significance of entanglement can be seen, for example, in the following quotation by Schrödinger [18],

“...I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives have become entangled. ...”

The astonishing features of entanglement were first brought to our attention in 1935 via the influential work by Einstein, Podolsky and Rosen (henceforth abbreviated as EPR) [20], and subsequently popularized by Schrödinger’s thought experiment on an innocent cat [19].¹ Specifically, in Ref. [20], EPR considered a pair of physical systems that are so strongly correlated that it becomes possible to predict, *with certainty*, some properties of the distant physical system by simply performing measurements on the local one. Exploiting such bizarre correlations offered by entanglement, EPR eventually came to the conclusion that the quantum mechanical predictions of *physical reality* is incomplete [20], just as statistical mechanics is incomplete within the framework of classical mechanics [22, 23].

For a long time after that, discussions arising out of EPR’s paper remained largely a philosophical debate. However, as Bell [5, 24] showed in the 1960s, the possibility of completing the quantum mechanical predictions in the way that EPR sought does lead

¹See also the English translation by Trimmer [21].

to experimentally falsifiable consequences. In particular, by considering a variant of EPR's argument due to Bohm [25], Bell [5] showed that quantum mechanical predictions on spatially separated systems cannot always be described by a *locally causal theory*. In such a theory, the outcomes of measurements are determined *in advance* merely by the choice of local measurements and some *local hidden variable* — which can be seen as information exchanged between the subsystems during their *common past*. Bell has thus ruled out the possibility of providing a locally causal description for all quantum phenomena — a brutal fact of life that is now succinctly called *Bell's theorem*.

Typically, the incompatibility between a locally causal description and the quantum mechanical prediction for a quantum state due to some choice of observables is revealed by the violation of *Bell inequalities*, which are statistical constraints that have to be satisfied by all locally causal theories. Since the early 1980s, there have been numerous experiments reporting Bell inequality violation in various physical systems (see, for example, Refs. [26, 27, 28]). While it is clear that entanglement is necessary to demonstrate a Bell inequality violation, by generalizing the notion of entanglement to mixed states, Werner [29] has found that not all entangled states can violate a Bell inequality (see also Refs. [30, 31, 32, 33]). In fact, it is not even known if all multipartite pure entangled states are Bell-inequality-violating [34, 35, 36, 37]. This state of affairs has inspired some to consider more general, nonstandard Bell experiments to reveal the bizarre correlations hidden in quantum states. In this regard, it was later shown by Popescu [38] and others [39, 40] that if a Bell experiment is preceded with appropriate *local preprocessing*, then a non-Bell-inequality-violating quantum state may become Bell-inequality-violating — a phenomenon that is now known as *hidden nonlocality*.

In recent years, the rising field of quantum information processing has also brought a resurgent interest in the study of Bell inequality violation. The pioneering work in this regard is due to Ekert [41], who showed that Bell inequality violation can be used to guarantee the security of a class of quantum key distribution protocols. Since then, a great deal of work has been carried out in this regard (see, for example, Refs. [42, 43, 44, 45, 46] and references therein). In fact, recently, it has even been argued in Refs. [45, 46] that Bell-inequality violation is necessary to guarantee the security of some entanglement-based quantum key distribution protocols. On the other hand, Bell inequality violation was also found to be relevant in other quantum information processing tasks, such as reduction of communication complexity [47, 48, 49]. In the context of quantum teleportation [50], Horodecki *et al.* [51] have shown that all two-qubit states violating a Bell inequality are useful for teleportation; Popescu, however, has shown that some two-qubit states not violating the same Bell inequality are also useful for teleportation [52]. Of course, given that quantum entanglement is an essential ingredient in many quantum information processing protocols [43], it is by no means accidental that a verification of entanglement through Bell inequality violation is carried out daily in many laboratories in the world.

Given the importance of Bell inequality violation, both from a foundational point of view and its relevance in quantum information processing, it is perhaps surprising that there are still many open problems related to the study of Bell inequality violation [53]. In particular, little is known as to which quantum states can violate a Bell inequality, both in a standard scenario and in a nonstandard scenario which also involves local preprocessing.

Even when a quantum state is known to violate a Bell inequality, the extent of violation is in most cases not well-quantified. On a related note, the maximal violation that quantum mechanics allows for a given Bell inequality is also not well-studied beyond some simple cases [17, 54, 55, 56, 57, 58].

The main goal of this thesis is to clarify the relationship between Bell inequality violation and quantum entanglement by determining the set of quantum states that can give rise to nonclassical behavior. The structure of this thesis is as follows. From Chapter 2 – Chapter 4, we will provide a comprehensive review of the theoretical background of the thesis. Specifically, Chapter 2 deals with some of the important concepts relevant to local causality and the key historical developments leading to Bell’s theorem. Then in Chapter 3, we will give a more technical introduction to the set of classical correlations,² which includes a formal introduction to the idea of a tight [59, 60], or facet-inducing Bell inequality [61]. Some of the well-known tight Bell inequalities will also be reviewed. After that, we will proceed to the quantum regime in Chapter 4 and introduce the notion of quantum correlation following Ref. [62]. Some well-known examples of entangled quantum states admitting a locally causal description will then be reviewed.

Most of our new research findings can be found in the second part of the thesis, from Chapter 5 – Chapter 7, while the rest are left in the appendices. In Chapter 5, we will present new findings in relation to the problem of determining if a given quantum state can violate some fixed but arbitrary Bell inequality via a standard Bell experiment. In particular, using convex optimization techniques [63] in the form of a semidefinite program [64], we have derived two algorithms to determine, respectively, a lower bound and an upper bound on the strength of correlation that a quantum state ρ can display in some given Bell experiments. These tools are also applied in Chapter 6 where we will look at some of the best known Bell inequality violations displayed by entangled states. Given that in practice, a Bell experiment involves measurements on many copies of the same quantum systems, we also investigated the possibility of getting a better Bell inequality violation by using collective measurements without postselection; this is the other subject of discussion in Chapter 6. Next, in Chapter 7, we will look into the possibility of deriving nonclassical correlations from all entangled quantum states. In particular, with the aid of an ancilla state which does not violate the Bell-CHSH inequality, we will provide a protocol to demonstrate a Bell-CHSH inequality violation coming from all bipartite entangled states. This provides a positive answer to the long-standing question of whether all bipartite entangled states can lead to some kind of observable nonclassical correlations. Finally, we will conclude with a summary of key results and some possibilities for future research in Chapter 8.

²This is the set of correlations allowed by a locally causal theory.

2

Bell's Theorem and Tests of Local Causality

In this chapter, we will give a brief historical review of the study of local causality in quantum mechanics. We will begin with the *incompleteness* arguments presented by Einstein, Podolsky and Rosen [20], and see how that had led to the celebrated discovery by Bell [5, 24]. After that, some of the key developments towards an experimental test of local causality will also be reviewed.

2.1 Bell's Theorem

2.1.1 The Einstein-Podolsky-Rosen Incompleteness Arguments

Quantum mechanics, as is well-known, only gives predictions, via the *wavefunction* or *state vector*, on the probabilities of obtaining a certain outcome in an experiment (see, for example, Ref. [65, 66]). Moreover, according to *Bohr's complementarity* [67, 68, 69], physical quantities described by two non-commuting observables in the theory are incompatible in that a complete knowledge of one precludes any knowledge of the other. This scenario is clearly in discord with the classical intuition that objective properties of physical systems exist independent of measurements.

Among those who were unsatisfied with Bohr's complementarity were Einstein, Podolsky and Rosen (EPR) who together put forward, in their 1935 paper [20], the argument that any *complete* physical theory must be such that

“every element of the physical reality must have a counterpart in the physical theory.”

A sufficient condition for the *reality* of a physical quantity that they have provided is as follows [20]:

“if without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.”

According to these criteria, Bohr's complementarity implies at least one of the followings, namely, (1) quantum mechanics is not a complete theory, or (2) the two physical quantities corresponding to non-commuting observables cannot have simultaneous reality. Moreover, by considering local measurements on two physical systems that have interacted in the past but are separated at the time of measurements, EPR came to the conclusion that if (1) is false, so is (2).

As an example, EPR considered a two-particle system described by the wavefunction

$$|\Psi(x_1, x_2)\rangle = \int_{-\infty}^{\infty} dp e^{(ip/\hbar)(x_1 - x_2 + x_0)}, \quad (2.1)$$

where x_1 and x_2 are, respectively, the coordinates attached to the two particles, x_0 is some arbitrary constant and p is the eigenvalue of the momentum operator for the first particle. It is not difficult to see that for both position and momentum measurements on the two particles, the outcomes derived are always perfectly correlated. In particular, if Alice and Bob are, respectively, at the receiving ends of the two particles, their measurement outcomes on these particles will read:

Measurement	Alice	Bob
Momentum P	p	$-p$
Position Q	x	$x + x_0$

Therefore, according to the criterion set up by EPR, should P be measured on the first particle, the momentum of the second particle is an element of physical reality; whereas if Q is measured on the first particle, the position of the second particle is an element of physical reality. Moreover [20],

since at the time of measurement the particles no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system.

Hence, by EPR's criterion of *reality*, both P and Q of the second particle, though corresponding to noncommuting observables in the theory, can have simultaneous reality, corresponding to the negation of (2). Since negation of (1) also led to the negation of (2), while at least one of (1) and (2) has to be true, EPR concluded that the quantum mechanical description of physical reality given by wavefunction is incomplete. Furthermore, at the very end of the paper [20], EPR optimistically expressed their belief that a theory that provides a *complete* description of physical reality is possible.

2.1.2 Completeness and Hidden-Variable Theory

Although no explicit proposal was given by EPR, it was commonly inferred from their arguments and the success of statistical mechanics that a complete description of physical reality can be attained if unknown (*hidden*) variables are supplemented to the wavefunction description of physical reality (see, for example Ref. [23] and references therein). Indeed, not known to EPR and many other founding fathers of quantum mechanics, towards the end of 1920s, de Broglie constructed a hidden-variable theory that is capable of explaining the quantum interference phenomena while retaining the corpuscular feature of individual particles [70, 71].

Despite that, the idea of completing the description given by quantum mechanics with additional variables has received much criticism over the years (see for example Ref. [24] and references therein). Among them, von Neumann's proof (pp. 305, Ref. [72]) of the impossibility of (*noncontextual*) hidden variables probably provided peace in mind to most of those who were against the proposal. The proof given by von Neumann in Ref. [72] has, nevertheless, imposed unnecessary restrictions on the unknown variables [24]. In fact, this was made blatant after Bohm rediscovered the hidden-variable theory [73, 74] first formulated by de Broglie [70, 71].

Nonetheless, *Bohmian mechanics* or the *pilot-wave model*, as the de Broglie-Bohm hidden-variable theory is currently known, was dismissed by many physicists because of the explicit “*nonlocal*” flavor in the theory. Ironically, it was precisely the discovery of this controversial theory that led Bell [24, 75] to consider, instead, the possibility of a *local hidden-variable theory* and hence his important discovery in 1964 [5].

For Einstein, he was firmly convinced that (pp 672, [22])

“... within the framework of future physics, quantum theory takes an analogous position as statistical mechanics takes within the framework of classical mechanics.”

Adhering to the same philosophy, Bell's consideration of a hidden-variable theory [5] is such that an average over some unknown ensemble labeled by the *hidden-variable* gives rise to the statistical behavior of quantum mechanical prediction. As Bell emphasized, the variables are *hidden* because they are not known to exist; they are not even accessible in principle, otherwise “quantum mechanics would be observably inadequate” [24, 76].

Clearly, not all hidden-variable theories are welcome in the physics community [24]. For instance, in the hidden-variable theory formulated by de Broglie and Bohm [73, 74], the trajectory of one particle may depend explicitly on the trajectory as well as the wavefunction of other particles that it has interacted with in the past, regardless of their spatial separation. This “*nonlocal*” feature of the theory is in apparent contradiction with the well-established intuition of causality that we have learned from special theory of relativity. Therefore, following EPR's flavor, Bell considered hidden-variable theories that are *local* such that, in Bell's words [5]:

“... the result of measurement on one system be unaffected by operations on a distant system with which it has interacted in the past ...”

In later years, a theory that satisfies Bell's notion of locality, or more specifically

“The direct causes (and effects) of events are near by, and even the indirect causes (and effects) are no further away than permitted by the velocity of light.”

is said to be *locally causal* [77]. Hereafter, we will use the term local hidden-variable theory (henceforth abbreviated as LHVT) and the term locally causal theory interchangeably.¹ As we shall see below, Bell's greatest contribution came in by showing that *quantum mechanics is not a locally causal theory* [5, 24].

2.1.3 Quantum Mechanics is not a Locally Causal Theory

To illustrate this remarkable fact of life, Bell [5] has chosen to work within the framework first presented by Bohm (Sec 15 – 19, Chap 22, Ref. [25]) concerning the spin degrees of freedom of two spin- $\frac{1}{2}$ particles, which is the analog of EPR's scenario for *discrete variable quantum systems*.² In this version of EPR's argument, pairs of spin- $\frac{1}{2}$ particles are prepared in the spin singlet state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_{\mathcal{A}} |\downarrow\rangle_{\mathcal{B}} - |\downarrow\rangle_{\mathcal{A}} |\uparrow\rangle_{\mathcal{B}}), \quad (2.2)$$

where $|\uparrow\rangle_{\mathcal{A}}$ and $|\downarrow\rangle_{\mathcal{A}}$ are correspondingly the *spin up* and *spin down* state of one of the particles with respect to some spatial direction³ (likewise for $|\uparrow\rangle_{\mathcal{B}}$ and $|\downarrow\rangle_{\mathcal{B}}$). After that, particles in each pair are separated and sent to two experimenters (hereafter always denoted by Alice and Bob), who can subsequently perform spin measurements along some (arbitrary) direction $\hat{\alpha}$ and $\hat{\beta}$, respectively, on these particles (c.f. Figure 2.1). Now, recall from quantum mechanics that the expectation value of such measurements reads

$$E_{\text{QM}}(\hat{\alpha}, \hat{\beta}) \equiv \langle \Psi^- | \sigma_{\hat{\alpha}} \otimes \sigma_{\hat{\beta}} | \Psi^- \rangle = -\hat{\alpha} \cdot \hat{\beta}, \quad (2.3)$$

where

$$\sigma_{\hat{\alpha}} \equiv \hat{\alpha} \cdot \vec{\sigma}, \quad \sigma_{\hat{\beta}} \equiv \hat{\beta} \cdot \vec{\sigma}, \quad (2.4)$$

$$\vec{\sigma} \equiv \sum_{l=x,y,z} \sigma_l \hat{e}_l, \quad (2.5)$$

\hat{e}_x is the unit vector pointing in the positive x direction (likewise for \hat{e}_y and \hat{e}_z) and

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.6)$$

are the Pauli matrices (here, we adopt the convention that $\sigma_z |\uparrow\rangle = |\uparrow\rangle$, $\sigma_z |\downarrow\rangle = -|\downarrow\rangle$). Thus, if $\hat{\alpha} = \hat{\beta}$, the measurement outcomes on both sides must be perfectly (anti-) correlated, i.e., if Alice's measurement outcome reads “ \uparrow ”, Bob's measurement outcome must read “ \downarrow ”.

¹The other terminology that is also commonly found in the literature is *local realistic theory*; this is however not as universally accepted, see e.g. Ref. [78].

²Incidentally, the experimental situation described in the original EPR paper [20] can indeed be explained within the framework of a locally casual theory [79].

³Since the spin singlet state is isotropic, the actual space quantization axis is immaterial.

Since this is true for other pair of $\hat{\alpha}'$ and $\hat{\beta}'$ such that $\hat{\alpha}' = \hat{\beta}'$, hence, by virtue of EPR's original argument, one can conclude that the “spin” along any direction for both of these particles must be “element of physical reality”.

Now, let us follow Ref. [5] and denote by λ any additional parameters carried by the particles that could provide a complete specification for these physical realities. Physically, we can think of λ as information that is exchanged between the particles during the preparation procedure but which is not completely encoded in the state vector $|\Psi^-\rangle$. As remarked in Ref. [5], the exact nature of λ is irrelevant, it could refer to a single or a set of random variables, or even a set of functions and it could take on continuous as well as discrete values. If we denote by o_a and o_b , respectively, the measurement outcome observed at Alice's and Bob's side, then by Bell's requirement of *locality*, we must have o_a as a function of λ and $\hat{\alpha}$ but not $\hat{\beta}$; likewise for o_b . Furthermore, the measurement outcome at each side is completely determined by these parameters such that [5]

$$o_a(\hat{\alpha}, \lambda) = \pm 1, \quad o_b(\hat{\beta}, \lambda) = \pm 1; \quad (2.7)$$

here, we adopt the convention that measurement outcomes “ \uparrow ” and “ \downarrow ” are assigned the value “+1” and “-1” respectively. Let us now define the *correlation function* as

$$E(\hat{\alpha}, \hat{\beta}) \equiv \int_{\Lambda} d\lambda \rho_{\lambda} o_a(\hat{\alpha}, \lambda) o_b(\hat{\beta}, \lambda), \quad (2.8)$$

where Λ is the space of hidden-variable and ρ_{λ} is some normalized probability density such that

$$\int_{\Lambda} d\lambda \rho_{\lambda} = 1. \quad (2.9)$$

Physically, the correlation function, Eq. (2.8), is just the average of the product of local measurement outcomes over an ensemble of physical systems characterized by some distribution of hidden-variable, ρ_{λ} . It then follows that a necessary condition for getting a *complete description* of the above-mentioned physical realities using local hidden-variable is that for all $\hat{\alpha}$ and $\hat{\beta}$

$$E(\hat{\alpha}, \hat{\beta}) = E_{\text{QM}}(\hat{\alpha}, \hat{\beta}) \quad (2.10)$$

for some choice of $o_a(\hat{\alpha}, \lambda)$, $o_b(\hat{\beta}, \lambda)$ and some choice of ρ_{λ} that is independent of $\hat{\alpha}$ and $\hat{\beta}$. As we shall see below, Eq. (2.10) cannot be made true in general. Nonetheless, it is interesting to note that Bell has constructed a specific *local hidden-variable model*⁴ (henceforth abbreviated as LHV) that makes it true for the case when $\hat{\alpha} \cdot \hat{\beta} = +1, 0, -1$ [5].

To show that Eq. (2.10) cannot be made true for all possible choices of measurement parameters, Bell introduced another unit vector $\hat{\beta}'$ and considered the following combination of correlation functions:

$$E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}') = \int_{\Lambda} d\lambda \rho_{\lambda} \left[o_a(\hat{\alpha}, \lambda) o_b(\hat{\beta}, \lambda) - o_a(\hat{\alpha}, \lambda) o_b(\hat{\beta}', \lambda) \right].$$

⁴Throughout this thesis, we will use the term *local hidden-variable model* to refer to, say, a set of rules, that can be used to reproduce some set of experimental statistics; it is less general than a LHVT, which is supposed to be able to reproduce all experimental statistics generated by quantum mechanics.

From triangle inequality, Eq. (2.7) and Eq. (2.9), it follows that

$$\begin{aligned} \left| E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}') \right| &\leq \int_{\Lambda} d\lambda \rho_{\lambda} \left| o_a(\hat{\alpha}, \lambda) o_b(\hat{\beta}, \lambda) \right| \left[1 - o_b(\hat{\beta}, \lambda) o_b(\hat{\beta}', \lambda) \right], \\ &= \int_{\Lambda} d\lambda \rho_{\lambda} \left[1 - o_b(\hat{\beta}, \lambda) o_b(\hat{\beta}', \lambda) \right], \\ &= 1 - \int_{\Lambda} d\lambda \rho_{\lambda} o_b(\hat{\beta}, \lambda) o_b(\hat{\beta}', \lambda). \end{aligned} \quad (2.11)$$

When $\hat{\alpha} = \hat{\beta}$, it follows from Eq. (2.3) that $E_{\text{QM}}(\hat{\alpha}, \hat{\beta}) = -1$. Therefore, Bell further assumed in Ref. [5] that if the measurement parameters chosen by both observers coincide, the outcomes of measurement, as determined by the hidden variables are also *perfectly correlated*:

$$o_a(\hat{\alpha}, \lambda) = -o_b(\hat{\alpha}, \lambda). \quad (2.12)$$

With this assumption, the above inequality becomes

$$\left| E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}') \right| - E(\hat{\beta}, \hat{\beta}') - 1 \leq 0, \quad (2.13)$$

which gives the very first inequality that has to be *satisfied by any LHVT* in the literature [5]. In the spirit of Bell's original work, let us introduce the following definition for a Bell inequality.⁵

Definition 1. *A Bell inequality is an inequality derived from the assumptions of a general local hidden-variable theory.*

In Ref. [5], Bell subsequently gave a formal proof, based on Eq. (2.13), that $E_{\text{QM}}(\hat{\alpha}, \hat{\beta})$ cannot equal or even be approximated arbitrarily closely by $E(\hat{\alpha}, \hat{\beta})$. However, to illustrate the point that quantum mechanics also gives rise to predictions not allowed by any LHVT, it suffices to show that for some choice of measurement parameters, the quantum mechanical version of Eq. (2.13), namely,

$$\left| E_{\text{QM}}(\hat{\alpha}, \hat{\beta}) - E_{\text{QM}}(\hat{\alpha}, \hat{\beta}') \right| - E_{\text{QM}}(\hat{\beta}, \hat{\beta}') - 1 \leq 0, \quad (2.14)$$

is violated. To this end, let us assume that all the spin measurements are performed on the $x - z$ plane and that $\hat{\alpha}$ points along the direction of the positive z -axis, i.e., $\hat{\alpha} = \hat{e}_z$. Then, for the choice of

$$\hat{\beta} = \frac{\sqrt{3}}{2} \hat{e}_x + \frac{1}{2} \hat{e}_z, \quad \hat{\beta}' = \frac{\sqrt{3}}{2} \hat{e}_x - \frac{1}{2} \hat{e}_z, \quad (2.15)$$

it can be easily verified using Eq. (2.3) that quantum mechanics predicts 1/2 for the *lhs* of inequality (2.14), thereby demonstrating that quantum mechanical prediction is, in general, incompatible with that given by any LHVT, c.f. Eq. (2.13).

The above finding gives rise to the following important theorem first derived by Bell [5]:

Theorem 2. *No local hidden-variable theory can reproduce all quantum mechanical predictions. Equivalently, quantum mechanics is not a locally causal theory.*

⁵It is worth noting that among the physics community, the term Bell inequality, or Bell-type inequality has sometimes been used to refer to inequality that arises out of an entanglement witness. To appreciate the distinction between these two kinds of inequalities, see, for example, Refs. [80, 81].

2.2 Towards an Experimental Test of Local Causality

2.2.1 Bell-Clauser-Horne-Shimony-Holt Inequality

The inequality (2.13) derived by Bell [5] has clearly demonstrated that *some* quantum mechanical predictions, in the ideal scenario, cannot be reproduced by any LHV. However, the assumption of *perfect correlation*, c.f. Eq. (2.12), or equivalently,

$$E(\hat{\alpha}', \hat{\beta}) = -1, \quad (2.16)$$

for $\hat{\alpha}' = \hat{\beta}$ is too strong to be justified in any realistic experimental scenario. The Bell inequality (2.13) was therefore not readily subjected to any experimental test. A few years later, in 1969, a resolution was provided by Clauser, Horne, Shimony and Holt (henceforth abbreviated as CHSH) who, instead of Eq. (2.16), assumed that for some $\hat{\alpha}'$ [6]

$$E(\hat{\alpha}', \hat{\beta}) = -1 + \delta, \quad (2.17)$$

where $0 \leq \delta \leq 1$. To conform with the prediction given by quantum mechanics, one expects that for spin measurement on the singlet state and when $\hat{\alpha}'$ is (approximately) aligned with $\hat{\beta}$, δ is close to but not exactly equal to zero.

Now, let's take this *imperfect correlation* into account by dividing the space of hidden-variable Λ into Λ_{\pm} such that

$$\Lambda_{\pm} = \{\lambda | o_a(\hat{\alpha}', \lambda) = \pm o_b(\hat{\beta}, \lambda)\}. \quad (2.18)$$

Then, it follows from Eq. (2.8), Eq. (2.9), Eq. (2.17) and Eq. (2.18) that

$$2 \int_{\Lambda^-} d\lambda \rho_{\lambda} = 2 - \delta. \quad (2.19)$$

Instead of inequality (2.13), inequality (2.11) now leads to

$$\begin{aligned} |E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}')| &\leq 1 - \int_{\Lambda^+} d\lambda \rho_{\lambda} o_b(\hat{\beta}, \lambda) o_b(\hat{\beta}', \lambda) - \int_{\Lambda^-} d\lambda \rho_{\lambda} o_b(\hat{\beta}, \lambda) o_b(\hat{\beta}', \lambda), \\ &= 1 - \int_{\Lambda^+} d\lambda \rho_{\lambda} o_a(\hat{\alpha}', \lambda) o_b(\hat{\beta}', \lambda) + \int_{\Lambda^-} d\lambda \rho_{\lambda} o_a(\hat{\alpha}', \lambda) o_b(\hat{\beta}', \lambda), \\ &= 1 - E(\hat{\alpha}', \hat{\beta}') + 2 \int_{\Lambda^-} d\lambda \rho_{\lambda} o_a(\hat{\alpha}', \lambda) o_b(\hat{\beta}', \lambda), \\ &\leq 1 - E(\hat{\alpha}', \hat{\beta}') + 2 \int_{\Lambda^-} d\lambda \rho_{\lambda} |o_a(\hat{\alpha}', \lambda) o_b(\hat{\beta}', \lambda)|, \\ &= 3 - E(\hat{\alpha}', \hat{\beta}') - \delta, \end{aligned}$$

which, together with Eq. (2.17), becomes

$$|E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}')| + E(\hat{\alpha}', \hat{\beta}) + E(\hat{\alpha}', \hat{\beta}') \leq 2. \quad (2.20)$$

This is the famous Bell-CHSH inequality that was first derived in Ref. [6]. It is interesting to note that a few years later [76], Bell gave an alternative derivation⁶ of inequality (2.20) by respectively replacing Eq. (2.7) and Eq. (2.8) with

$$|\bar{o}_a(\hat{\alpha}, \lambda)| \leq 1, \quad |\bar{o}_b(\hat{\beta}, \lambda)| \leq 1, \quad (2.22)$$

and

$$E(\hat{\alpha}, \hat{\beta}) \equiv \int_{\Lambda} d\lambda \, \rho_{\lambda} \, \bar{o}_a(\hat{\alpha}, \lambda) \, \bar{o}_b(\hat{\beta}, \lambda). \quad (2.23)$$

Here, Bell tried to be more general (as compared with his approach in Ref. [5]) by assuming that the measurement apparatuses could also contain hidden-variable that could influence the experimental results. In the above expressions, $\bar{o}_a(\hat{\alpha}, \lambda)$ is thus used to denote an average over the hidden-variable associated with Alice's apparatus when it is set to perform measurements parameterized by $\hat{\alpha}$; similarly for $\bar{o}_b(\hat{\beta}, \lambda)$.

At this stage, it is worth making a few other remarks. Firstly, in contrast with Bell's first inequality, Eq. (2.13), that was developed for spin measurements on the singlet state, the Bell-CHSH inequality is also relevant to other physical states as well as other physical systems. In fact, it is applicable, as a constraint imposed by LHVTs, to *any* experimental statistics involving *two* spatially separated subsystems and where *two dichotomic*⁷ measurements — each giving outcomes labeled by ± 1 — can be performed on each of the subsystems. Essentially, this means that in the more general experimental framework, the parameters $\hat{\alpha}$ etc. are merely labels to distinguish the different measurements that Alice and Bob may perform on the subsystem in their possession.

As a result, and for the convenience of subsequent discussion, let us introduce the following notation for the correlation function associated with Alice measuring the observable A_{s_a} and Bob measuring the observable B_{s_b} , i.e.,

$$E(A_{s_a}, B_{s_b}) \equiv \int_{\Lambda} d\lambda \, \rho_{\lambda} \, o_a(A_{s_a}, \lambda) \, o_b(B_{s_b}, \lambda), \quad (2.24)$$

where the outcomes of local measurements o_a and o_b are now functions of the hidden variable λ and, respectively, the local observables A_{s_a} and B_{s_b} . In particular, if we now make the following associations between the measurement parameters $\{\alpha, \alpha', \beta, \beta'\}$ and the local observables $\{A_{s_a}, B_{s_b}\}_{s_a, s_b=1}^2$:

$$\hat{\alpha} \rightarrow A_2, \quad \hat{\alpha}' \rightarrow A_1, \quad \hat{\beta} \rightarrow B_1, \quad \hat{\beta}' \rightarrow B_2, \quad (2.25)$$

it is clear that both inequality (2.20) and inequality (2.21) imply the following inequality:

$$E(A_1, B_1) + E(A_1, B_2) + E(A_2, B_1) - E(A_2, B_2) \leq 2. \quad (2.26)$$

⁶Strictly, the inequality that was later derived by Bell reads:

$$\left| E(\hat{\alpha}, \hat{\beta}) - E(\hat{\alpha}, \hat{\beta}') \right| + \left| E(\hat{\alpha}', \hat{\beta}) + E(\hat{\alpha}', \hat{\beta}') \right| \leq 2, \quad (2.21)$$

but as we shall see below, we can essentially treat it as the same inequality as that given by Eq. (2.20).

⁷A dichotomic measurement is one that yields one out of two possible outcomes.

Evidently, if this is a valid constraint that has to be satisfied by any LHVT, so is any other obtained by relabeling the local observers (“Alice” \leftrightarrow “Bob”), local measurement settings ($A_1 \leftrightarrow A_2$, $B_1 \leftrightarrow B_2$) and/or outcomes ($+1 \leftrightarrow -1$). For example, if we instead make the associations $\hat{\alpha} \rightarrow A_1$, $\hat{\alpha}' \rightarrow A_2$ and relabel all the $+1$ outcomes at Alice’s site by -1 and *vice versa*, then we will arrive at

$$-2 \leq E(A_1, B_1) - E(A_1, B_2) + E(A_2, B_1) + E(A_2, B_2), \quad (2.27)$$

which is clearly different from inequality (2.26). Nonetheless, the difference between these inequalities, which is due to a different choice of labels, is physically irrelevant. After all, when testing a set of experimental data against a Bell inequality, the choice of these labels is completely arbitrary. As such, let us define the equivalence class of Bell inequalities as follows [59, 60].

Definition 3. *A Bell inequality is equivalent to another if and only if one can be obtained from the other by relabeling the local observers, local measurement settings and/or measurement outcomes.*

Under this definition, it is straightforward to see that apart from inequality (2.27), inequality (2.26) is also equivalent to 6 other inequalities. Hereafter, when there is no risk of confusion, we will refer to inequality (2.26) as the Bell-CHSH inequality and to the entire class of 8 inequalities that are equivalent to inequality (2.26) as the Bell-CHSH inequalities. In relation to inequality (2.20), it is also not difficult to see that this inequality is violated if and only if (at least) one of the Bell-CHSH inequalities is violated; likewise for inequality (2.21).

As a last remark, we note that the Bell-CHSH inequality is an example of what is now called a *(Bell) correlation inequality* — a Bell inequality that only involves linear combination of *correlation functions*. Clearly, a correlation function, which can be determined experimentally by averaging over the product of the outcome of local observables, is not the only quantity that is derivable from a given set of experimental data; the relative frequency of experimental outcomes, in the limit of large sample size, gives a good approximation to the probability of obtaining that particular outcome. In the next section, we will look at an example of the other prototype of (linear) Bell inequalities, namely, one that involves a linear combination of joint and marginal probabilities of experimental outcomes.

2.2.2 Bell-Clauser-Horne Inequality

The earlier work by CHSH is no doubt a big step towards an experimental test for the feasibility of locally casual theories. However, due to imperfect detection and other realistic experimental concerns, the Bell-CHSH inequality (2.26) can only be put into a real experimental test when supplemented with an auxiliary assumption on the ensemble of detected particles [6, 82]. Specifically, in the context of polarization measurement on photons, the original assumption made by CHSH is that if a pair of photons emerges from the respective polarizers located at Alice’s and Bob’s side, the probability of their joint detection is independent of the orientation of the polarizers.

A few years later, work by Clauser and Horne (hereafter abbreviated as CH) demonstrated that without an auxiliary assumption, neither the experiment carried out by Freedman and Clauser [83] nor any similar ones with improved detector efficiency can give a definitive test of locally causal theories [7]. To remedy the problem, CH derived, in the same paper [7], another Bell inequality and showed that when supplemented with a considerably weaker *no enhancement* assumption, the results obtained by Freedman and Clauser are indeed incompatible with LHVTs [7].

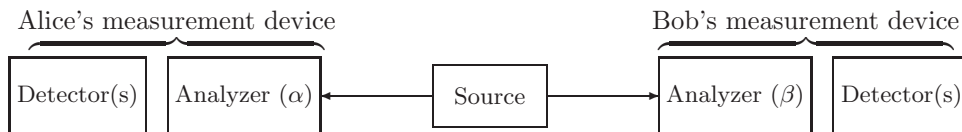


FIGURE 2.1: Schematic diagram of the experimental setup involved in a standard two-party Bell experiment. The source produces pairs of physical systems that are subsequently distributed, respectively, to Alice and Bob. They then subject the physical system that they receive to an analyzer which has an adjustable parameter (denoted by α and β correspondingly). For example, in the case of polarization measurement on photons, an analyzer is simply a combination of waveplates and a polarizer. The final stage of the measurement process consists of detecting the subsystems that pass through each analyzer with one or more detectors. In the scenario considered by Bell [5] and Clauser *et al.* [6], there are two detectors at each site, whereas in the original experimental scenario considered by CH [7], there is only one detector after each analyzer.

The scenario that CH considered is a familiar one, namely, one that involves ensembles of two particles being sent to Alice and Bob respectively. Under the control of each experimenter is an analyzer with an adjustable parameter (denoted by α and β respectively) and a detector. At each run of the experiment, let us denote by λ the state of the two-particle system and $p_{AB}(\alpha, \beta, \lambda)$ the probability that for this two-particle state, a count is triggered at both detectors conditioned on Alice setting her analyzer to α and Bob setting his to β ; the marginal probabilities of detecting a particle $p_A(\alpha, \lambda)$ and $p_B(\beta, \lambda)$ are similarly defined. In these terminologies, the no enhancement assumption states that for a given state λ , the probability of detecting a particle with the analyzer removed is greater than or equal to the probability of detecting a particle when the analyzer is in place.

Now, note that for a given (normalized) probability density ρ_λ characterizing the ensemble of states emitted, the observed relative frequencies should correspond to

$$\begin{aligned} p_A(\alpha) &= \int_{\Lambda} d\lambda \rho_\lambda p_A(\alpha, \lambda), & p_B(\beta) &= \int_{\Lambda} d\lambda \rho_\lambda p_B(\beta, \lambda), \\ p_{AB}(\alpha, \beta) &= \int_{\Lambda} d\lambda \rho_\lambda p_{AB}(\alpha, \beta, \lambda). \end{aligned} \quad (2.28)$$

It is worth noting that as it is, the above formulation could very well be applied to quantum mechanical prediction, with the wavefunction $|\psi\rangle$ playing the role of λ . As with the correlation function, Eq. (2.8), the condition of local causality comes in by demanding that the probability of joint detection factorizes [7], i.e.,

$$p_{AB}(\alpha, \beta, \lambda) = p_A(\alpha, \lambda) p_B(\beta, \lambda). \quad (2.29)$$

From the definition of probabilities, it follows that

$$\begin{aligned} 0 \leq p_{\mathcal{A}}(\alpha, \lambda) \leq 1, \quad 0 \leq p_{\mathcal{A}}(\alpha', \lambda) \leq 1, \\ 0 \leq p_{\mathcal{B}}(\beta, \lambda) \leq 1, \quad 0 \leq p_{\mathcal{B}}(\beta', \lambda) \leq 1, \end{aligned} \quad (2.30)$$

where α' and β' are some other choice of parameters for the analyzers. Together, Eq. (2.29) and Eq. (2.30) imply that [7]

$$\begin{aligned} -1 \leq & p_{\mathcal{A}}(\alpha, \lambda) p_{\mathcal{B}}(\beta, \lambda) + p_{\mathcal{A}}(\alpha, \lambda) p_{\mathcal{B}}(\beta', \lambda) + p_{\mathcal{A}}(\alpha', \lambda) p_{\mathcal{B}}(\beta, \lambda) \\ & - p_{\mathcal{A}}(\alpha', \lambda) p_{\mathcal{B}}(\beta', \lambda) - p_{\mathcal{A}}(\alpha, \lambda) - p_{\mathcal{B}}(\beta, \lambda) \leq 0 \end{aligned}$$

for each given λ . After averaging over the ensemble space Λ , one arrives at

$$-1 \leq p_{\mathcal{AB}}(\alpha, \beta) + p_{\mathcal{AB}}(\alpha, \beta') + p_{\mathcal{AB}}(\alpha', \beta) - p_{\mathcal{AB}}(\alpha', \beta') - p_{\mathcal{A}}(\alpha) - p_{\mathcal{B}}(\beta) \leq 0, \quad (2.31)$$

which is the Bell-CH inequality — the very first Bell inequality for *probabilities* derived in the literature. Notice that to arrive at the lower limit of inequality (2.31), we also have to assume that the probability density ρ_{λ} is normalized, Eq. (2.9).

Let us now make a few other remarks concerning inequality (2.31). To begin with, we note that although the inequality was derived by considering a one-output-channel analyzer that is followed by a single detector, it could very well be applied to measurement devices equipped with two (or more) detectors, thereby giving rise to two (or more) possible outcomes.⁸ In particular, for the specific case of two possible outcomes, which we will label as “ \pm ”, the same analysis allows us to arrive at the inequality [7]

$$p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{A}}^{o_a}(1) - p_{\mathcal{B}}^{o_b}(1) \leq 0, \quad (2.32a)$$

and

$$- [p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{A}}^{o_a}(1) - p_{\mathcal{B}}^{o_b}(1)] \leq 1, \quad (2.32b)$$

where each measurement outcome o_a and o_b can be “ \pm ” and $p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b)$ is now the probability of Alice observing outcome o_a and Bob observing outcome o_b conditioned on her performing the s_a^{th} measurement and him performing the s_b^{th} measurement; the marginal probabilities $p_{\mathcal{A}}^{o_a}(s_a)$ and $p_{\mathcal{B}}^{o_b}(s_b)$ are analogously defined. Notice that the four inequalities (2.32b) are actually equivalent to inequalities (2.32a) and can be obtained from the latter, for example, via the identity $p_{\mathcal{AB}}^{++}(s_a, s_b) + p_{\mathcal{AB}}^{+-}(s_a, s_b) = p_{\mathcal{A}}^{+}(s_a)$.

Let us also remark that the set of 8 inequalities given in Eq. (2.32) are symmetrical with respect to swapping \mathcal{A} & \mathcal{B} and have taken into account all possible ways of labeling of the outcomes. Nevertheless, additional equivalent inequalities, such as

$$-1 \leq p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{A}}^{o_a}(1) - p_{\mathcal{B}}^{o_b}(2) \leq 0 \quad (2.33)$$

can still be obtained by relabeling the local measurement settings. Hereafter, unless stated otherwise, the term Bell-CH inequality would refer to Eq. (2.32a) with only two possible outcomes.

⁸Strictly, there are three possible outcomes when there are two detectors, with the other possible outcome corresponding to no detection.

In relation to the Bell-CHSH inequality, we recall that the correlation function defined in Eq. (2.24) can actually be rewritten as⁹

$$E(A_{s_a}, B_{s_b}) = p_{\mathcal{AB}}^{++}(s_a, s_b) + p_{\mathcal{AB}}^{--}(s_a, s_b) - p_{\mathcal{AB}}^{+-}(s_a, s_b) - p_{\mathcal{AB}}^{-+}(s_a, s_b), \quad (2.34)$$

i.e., the average value of the product of observables or

$$E(A_{s_a}, B_{s_b}) = p_{\mathcal{AB}}^{o_a=o_b}(s_a, s_b) - p_{\mathcal{AB}}^{o_a \neq o_b}(s_a, s_b), \quad (2.35)$$

which is the difference between the probability of observing the same outcomes at the two sides and the probability of observing different outcomes at the two sides. Thus, by adding the two inequalities in Eq. (2.32a) with $o_a \neq o_b$ and subtracting them from the two inequalities with $o_a = o_b$, one arrives at the Bell-CHSH inequality in the form of Eq. (2.26). Conversely, if there are only two possible outcomes such that

$$p_{\mathcal{A}}^{+}(s_a) + p_{\mathcal{A}}^{-}(s_a) = 1 \quad \forall s_a, \quad p_{\mathcal{B}}^{+}(s_b) + p_{\mathcal{B}}^{-}(s_b) = 1 \quad \forall s_b, \quad (2.36)$$

then all the four Bell-CH inequalities given in Eq. (2.32a) can also be obtained from the Bell-CHSH inequalities via Eq. (2.34) or Eq. (2.35). Hence, when seen as a set of constraints imposed by LHVTs on two particles, where each of them is subjected to two alternative dichotomic measurements, the Bell-CH inequalities are entirely equivalent to the Bell-CHSH inequalities [7].

2.2.3 Experimental Progress

Since the late 1960s, many experiments have been carried out, via the Bell-CH and Bell-CHSH inequalities, to probe the adequacy of locally causal theories. An account of the early attempts prior to the 1980s can be found in the excellent review by Clauser and Shimony [82]. These early results, however, were not compelling enough to close the debate due to the various possible loopholes in experiments [84].

Among which, the *communication loophole* survived happily till the influential experiment performed by Aspect and coworkers in 1982 using time-varying analyzers [85]. Since then, many have considered the impossibility of a LHVT verified, even though some still think otherwise (see for example [86, 87, 88, 89, 90] and references therein). As of now, the experiment that most convincingly evades the communication loophole was carried out by Weihs and collaborators in 1998 [91]. The equally notorious *detection loophole* has also been closed quite recently by Rowe and coworkers [27]. A single experiment that closes both these loopholes at once is, nevertheless, still being sought [26, 92]. In this regard, it is worth noting that some other loopholes such as those considered in Refs. [93, 94] exist, but they are generally considered less compelling. For further information on recent Bell experiments, see the review by Genovese [95].

⁹To this end, we are identifying the s_a^{th} measurement at Alice's site as a measurement of A_{s_a} while the s_b^{th} measurement at Bob's site as a measurement of B_{s_b} .

3

Classical Correlations and Bell Inequalities

In the last chapter, we have seen two important examples of Bell inequalities that were developed in the hope of realizing a convincing test of local causality. Bell inequalities, nevertheless, can also be understood from a completely different perspective. Specifically, in this chapter, we will see that in the space of probability vectors, which we will call the *space of correlations*, the tight Bell inequalities correspond to hyperplanes that together form the boundaries of the *convex set* of classical correlations.¹ Froissart is apparently the pioneer of such a geometrical approach to Bell inequalities [96]. Not too long after that, this approach was discovered independently by Garg and Mermin [97]. A few years later, a general study along the same lines was also carried out by Pitowsky [62, 98]. A great advantage of this geometrical approach is that it can be easily generalized to more complicated experimental scenarios and hence, allows more complicated Bell inequalities to be derived in a systematic manner.

3.1 Classical Correlations and Probabilities

Before we move on to the more general scenario, let us first go through the following example of a hypothetical Bell experiment to gain some intuition. In particular, let us consider an experimental scenario where the Bell-CHSH inequality, or equivalently the two-outcome Bell-CH inequality, is applicable (Figure 2.1). Now, let us imagine that the experimental data collected (Table 3.1) — including those not explicitly shown in the table — satisfy the

¹Although our treatment focuses (almost) exclusively on probability vectors, it should be clear that one can just as well consider a space of correlations that is defined in terms of various correlation functions, as in Eq. (2.24). In that case, a (tight) Bell correlation inequality similarly defines a closed halfspace where the convex set of classical correlations resides.

following joint probabilities

$$p_{\mathcal{AB}}^{++}(1, 1) = 1, \quad p_{\mathcal{AB}}^{+-}(1, 1) = 0, \quad p_{\mathcal{AB}}^{-+}(1, 1) = 0, \quad p_{\mathcal{AB}}^{--}(1, 1) = 0, \quad (3.1a)$$

$$p_{\mathcal{AB}}^{++}(1, 2) = \frac{1}{2}, \quad p_{\mathcal{AB}}^{+-}(1, 2) = \frac{1}{2}, \quad p_{\mathcal{AB}}^{-+}(1, 2) = 0, \quad p_{\mathcal{AB}}^{--}(1, 2) = 0, \quad (3.1b)$$

$$p_{\mathcal{AB}}^{++}(2, 1) = \frac{1}{2}, \quad p_{\mathcal{AB}}^{+-}(2, 1) = 0, \quad p_{\mathcal{AB}}^{-+}(2, 1) = \frac{1}{2}, \quad p_{\mathcal{AB}}^{--}(2, 1) = 0, \quad (3.1c)$$

$$p_{\mathcal{AB}}^{++}(2, 2) = \frac{1}{4}, \quad p_{\mathcal{AB}}^{+-}(2, 2) = \frac{1}{4}, \quad p_{\mathcal{AB}}^{-+}(2, 2) = \frac{1}{4}, \quad p_{\mathcal{AB}}^{--}(2, 2) = \frac{1}{4}, \quad (3.1d)$$

and marginal probabilities

$$p_{\mathcal{A}}^{+}(1) = 1, \quad p_{\mathcal{A}}^{-}(1) = 0, \quad p_{\mathcal{A}}^{+}(2) = \frac{1}{2}, \quad p_{\mathcal{A}}^{-}(2) = \frac{1}{2}, \quad (3.1e)$$

$$p_{\mathcal{B}}^{+}(1) = 1, \quad p_{\mathcal{B}}^{-}(1) = 0, \quad p_{\mathcal{B}}^{+}(2) = \frac{1}{2}, \quad p_{\mathcal{B}}^{-}(2) = \frac{1}{2}. \quad (3.1f)$$

Evidently, we can collect all the 16 joint probabilities together and think of them as the components of a probability vector \mathbf{p} living in a 16-dimensional space. Let us now make the following definitions in relation to such a probability vector.

Definition 4. *A probability vector is said to be classical if it can be generated from some local hidden-variable model.*

Hereafter, we will also loosely refer to a probability vector as a correlation. This can be justified by noting that from the components of a probability vector, we can learn the extent to which measurement outcomes between subsystems \mathcal{A} and \mathcal{B} are correlated. For example, if \mathcal{A} and \mathcal{B} involved in the experiment are totally *uncorrelated*, we will expect that all the joint probabilities factorize and equal to the product of the corresponding marginal probabilities, i.e.,

$$p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b) = p_{\mathcal{A}}^{o_a}(s_a) p_{\mathcal{B}}^{o_b}(s_b). \quad (3.2)$$

Moreover, all experimental statistics that could be of interest, such as the correlation function for given measurement settings, as well as other higher order moments can be computed according to the standard procedures.

Now, let us again look at the set of experimental data presented in Table 3.1. If there exists a LHVM that can reproduce this set of data, we will be able to fill in the blanks corresponding to unperformed measurement results such that all the joint and marginal probabilities are preserved. Therefore, if the unfilled entries in the table can be filled up in such a way that respects all the probabilities listed in Eq. (3.1), we will have got a LHVM that reproduces all the experimental statistics, and hence correlations derivable from Table 3.1. An example of how this can be done is shown in Table 3.2. In this case, we can see n as an index for the local hidden-variable λ that is associated with each run of the experiment. Then, in each run n , once the choice of local measurement is decided, the outcome of the measurement can be read off directly from the table (regardless of the other entries listed in the same row of the table).

That a LHVM can be constructed for the data presented in Table. 3.1 is not incidental. Simple calculations using Eq. (2.32) and Eq. (3.1) show that none of the Bell-CH inequalities

Table 3.1: A hypothetical set of experimental data gathered in an experiment to test the Bell-CHSH inequality or the Bell-CH inequality. Here, n is an index to label each run of the experiment and N is some very large number such that the data set is statistically significant. The local measurements that may be performed by Alice are labeled by A_1 and A_2 whereas that for Bob are labeled by B_1 and B_2 . Outcomes of the experiments are labeled by ± 1 and are tabulated under the respective local measurements that are carried out in each run of the experiment.

n	A_1	A_2	B_1	B_2
1	1		1	
2	1			1
3		1	1	
4	1		1	
5		-1		1
\vdots	\vdots	\vdots	\vdots	\vdots
1000	1			-1

n	A_1	A_2	B_1	B_2
1001		-1		-1
1002	1		1	
1003		1		1
1004	1		1	
1005		-1	1	
\vdots	\vdots	\vdots	\vdots	\vdots
N		1		-1

Table 3.2: The same set of experimental data as in Table 3.1 but with the *unperformed* measurement results (enclosed within round brackets) filled in according to some hypothetical LHV. In particular, the LHV works in such a way that the newly filled entries in the table give rise to the same joint and marginal probabilities as the original entries listed in Table. 3.1, c.f. Eq. (3.1).

n	A_1	A_2	B_1	B_2
1	1	(1)	1	(1)
2	1	(-1)	(1)	1
3	(1)	1	1	(1)
4	1	(-1)	1	(-1)
5	(1)	-1	(1)	1
\vdots	\vdots	\vdots	\vdots	\vdots
1000	1	(-1)	(1)	-1

n	A_1	A_2	B_1	B_2
1001	(1)	-1	(1)	-1
1002	1	(1)	1	(-1)
1003	(1)	1	(1)	1
1004	1	(1)	1	(-1)
1005	(1)	-1	1	(1)
\vdots	\vdots	\vdots	\vdots	\vdots
N	(1)	1	(1)	-1

is violated by the experimental data presented in Table. 3.1. Evidently, no-violation of the Bell-CH inequality is a necessary condition for the existence of a LHV for the given experimental data. Nevertheless, as was first shown by Fine in 1982 [99], fulfillment of all the Bell-CH inequalities is also sufficient to guarantee the existence of a LHV, provided that the experimental data only involves two dichotomic measurements performed by two observers [100, 101, 102]. Hence, in an experimental scenario involving only two observers and two dichotomic measurements per site, a complete characterization of classical correlations can be obtained *solely* using the Bell-CH inequalities. What about experiments involving more observers, more local measurements per site, or more outcomes per measurement? These are the questions that we will discuss in the following sections.

3.2 Geometrical Structure of the Set of Classical Correlations

3.2.1 The Spaces of Correlations

For the subsequent discussion, let us consider a more general scenario whereby a source — characterized by some physical state ρ — distributes pairs of physical systems to Alice and Bob, and where each of them can perform (on the subsystems that they receive), respectively, m_A and m_B alternative measurements that would each generate n_A and n_B distinct outcomes.² For now, we will restrict our attention to this bipartite scenario, but most of the following arguments can be modified easily to cater for the multipartite scenario. In view of the forthcoming discussion, let us also introduce the vectors

$$\mathbf{m} \equiv (m_A, m_B), \quad \mathbf{n} \equiv (n_A, n_B) \quad (3.3)$$

for, respectively, a compact description of the number of local measurement settings and the number of possible outcomes for each local measurement. As with the previous section, the experimental statistics in such a scenario can be summarized as a probability vector $\mathbf{p} \in \mathbb{R}^{d_p}$ where $d_p = m_A m_B n_A n_B$ (if one prefers to work in the space of correlations that is defined only in terms of *full* correlation functions $E(A_{s_a}, B_{s_b})$, then we will be working in a space of dimension $m_A m_B$ — Sec. 3.3.2). The components of the probability vectors are the joint probabilities $p_{AB}^{o_a o_b}(s_a, s_b)$.³ We will refer to this real vector space as the *space of correlations*, denoted by $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$. Clearly, for our purpose, not all of the d_p coordinates in $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$ are independent. For instance, given a particular choice of Alice's and Bob's measurement, there must be an outcome at Alice's as well as Bob's site.⁴ Normalization of probability therefore requires:

$$\sum_{o_a=1}^{n_A} \sum_{o_b=1}^{n_B} p_{AB}^{o_a o_b}(s_a, s_b) = 1 \quad \forall s_a, s_b, \quad (3.4)$$

where we have labeled the outcomes registered at Alice's site as $o_a = 1, 2, \dots, n_A$ (likewise $o_b = 1, 2, \dots, n_B$ at Bob's site). Moreover, adhering to the principles of relativity, we shall be contented with correlations that do not allow faster-than-light signaling. These are correlations that respect the following equalities [103]:

$$\sum_{o_a=1}^{n_A} p_{AB}^{o_a o_b}(s_a, s_b) = p_B^{o_b}(s_b) \quad \text{and} \quad \sum_{o_b=1}^{n_B} p_{AB}^{o_a o_b}(s_a, s_b) = p_A^{o_a}(s_a) \quad \forall s_a, s_b. \quad (3.5)$$

In words, this means that the marginal probability of Alice observing local measurement outcome o_a , conditioned on her measuring s_a , i.e., $p_A^{o_a}(s_a)$ is independent of the choice of

²Of course, one can be more general than this and allows each measurement to have different number of possible outcomes. Nevertheless, for brevity, we shall be contented with a discussion on the case where all local measurements performed by Alice yield the same number of possible outcomes (likewise for Bob).

³One can, instead, work in a space of probabilities with dimension $d > d_p$ such that each probability vector \mathbf{p} also has the marginal probabilities $p_A^{o_a}(s_a)$ and $p_B^{o_b}(s_b)$ as components. However, this is not necessary, as the marginal probabilities are not independent from the joint probabilities.

⁴For the purpose of present discussion, one could treat the possibility of no-detection as one of the possible outcomes.

measurement s_b made by the spatially separated observer Bob; likewise for $p_B^{ob}(s_b)$. This is now commonly known as the *no-signaling condition* (see, for example, Refs. [104, 105]), which was originally termed the *relativistic causality condition* in Ref. [103].

By a simple counting argument, one can show that after taking into account all of these constraints, there are effectively only [60]

$$d'_p = m_A m_B (n_A - 1)(n_B - 1) + m_A (n_A - 1) + m_B (n_B - 1) \quad (3.6)$$

independent entries in the probability vector, which can be taken to be all but one of the marginal probabilities $p_A^{oa}(s_a)$ for each s_a , likewise for $p_B^{ob}(s_b)$, plus $(n_A - 1)(n_B - 1)$ of the joint probabilities $p_{AB}^{oaob}(s_a, s_b)$ for each combination of s_a and s_b . Hence, we are essentially only interested in a subspace of the set of probability vectors that is of dimension d'_p .

3.2.2 The Convex Set of Classical Correlations

Now, let us take a closer look at the set of classical correlations associated with the experimental scenario described above. Hereafter, we will denote this set by $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ (analogously, we will denote the set of classical correlations defined in the space of correlation functions as ${}^c\mathcal{P}_{m_A; m_B}^{n_A; n_B}$). From Eq. (2.28) and Eq. (2.29), it follows that a classical probability vector is one whose entries satisfy

$$p_{AB}^{oaob}(s_a, s_b) = \int_{\Lambda} d\lambda \, \rho_{\lambda} p_A^{oa}(s_a, \lambda) p_B^{ob}(s_b, \lambda) \quad (3.7)$$

for some choice of $p_A^{oa}(s_a, \lambda)$ and $p_B^{ob}(s_b, \lambda)$, and some probability density ρ_{λ} . For any two classical probability vectors \mathbf{p}_{LHV} and \mathbf{p}'_{LHV} , any convex combination of them gives rise to a probability vector

$$\mathbf{p}'' \equiv q \mathbf{p}_{\text{LHV}} + (1 - q) \mathbf{p}'_{\text{LHV}}, \quad (3.8)$$

that is also classical. This is because the resulting probability vector \mathbf{p}'' can be realized via a LHVM which consists of implementing the LHVM associated with \mathbf{p}_{LHV} and \mathbf{p}'_{LHV} stochastically. Specifically, by tossing a biased coin with probability q of getting heads and probability $1 - q$ of getting tails, the probability vector \mathbf{p}'' can be realized by implementing the LHVM associated with \mathbf{p}_{LHV} whenever the outcome of the toss is heads, and the LHVM associated with \mathbf{p}'_{LHV} whenever the outcome of the toss is tails. Therefore, the set of classical probability vectors $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ is *convex*.

A natural question that follows is: what are the extreme points⁵ of this set? With some thought, it is not difficult to see that probability vectors such that the joint probability factorizes, i.e.,

$$p_{AB}^{oaob}(s_a, s_b) = p_A^{oa}(s_a) p_B^{ob}(s_b), \quad (3.9a)$$

and for which the marginal probabilities are either 0 or 1, i.e.,

$$p_A^{oa}(s_a) = 0, 1, \quad p_B^{ob}(s_b) = 0, 1, \quad (3.9b)$$

⁵An extreme point of a convex set is a point in the set which cannot be expressed as a nontrivial convex combination of two or more different points in the set [106, 107].

are extreme points of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ [60]. These are probability vectors corresponding to deterministic LHVMs. Physically, each of these probability vectors corresponds to a scenario where the experimental outcomes for given local measurement settings are deterministic; once the local measurement setting is chosen, one and only one of the local detectors will ever click.⁶ Conversely, it is also not difficult to see from Eq. (3.7) and Eq. (3.9) that *any* other classical probability vectors can be written as a nontrivial convex combination of these extremal probability vectors. In other words, a probability vector is an extreme point of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ if and only if it satisfies Eq. (3.9).

Given that the physical scenario corresponding to an extreme point of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ is such that the local measurement settings determine the local measurement outcome with certainty, we might as well label each of these extreme points by two sets of indices \mathbf{a} and \mathbf{b} that are, respectively, associated with the measurement outcomes observed by Alice and Bob [108] (Figure 3.1). Specifically, let us denote by ${}^{\mathbf{a}, \mathbf{b}}\mathbf{B}_{AB}$ an extreme point of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, $\vartheta_{s_a}^{[1]}$ Alice's measurement outcome conditioned on her measuring A_{s_a} and $\vartheta_{s_b}^{[2]}$ Bob's measurement outcome conditioned on him measuring B_{s_b} . Then the two sets of indices $\mathbf{a} = (\vartheta_1^{[1]}, \vartheta_2^{[1]}, \dots, \vartheta_{m_A}^{[1]})$ where $\vartheta_{s_a}^{[1]} = 1, 2, \dots, n_A$ and $\mathbf{b} = (\vartheta_1^{[2]}, \vartheta_2^{[2]}, \dots, \vartheta_{m_B}^{[2]})$ where $\vartheta_{s_b}^{[2]} = 1, 2, \dots, n_B$ will completely characterize ${}^{\mathbf{a}, \mathbf{b}}\mathbf{B}_{AB}$ in the sense that its component reads [108]

$${}^{\mathbf{a}, \mathbf{b}}\mathbf{B}_{AB}^{o_a o_b}(s_a, s_b) = \delta_{o_a \vartheta_{s_a}^{[1]}} \delta_{o_b \vartheta_{s_b}^{[2]}}. \quad (3.10)$$

For finite number of local measurement settings and measurement outcomes, i.e.,

$$m_A, m_B, n_A, n_B < \infty,$$

it is possible to enumerate all of these extreme points by going through all legitimate boolean values of the local probabilities. In total, there are thus

$$n_v = n_A^{m_A} n_B^{m_B} \quad (3.11)$$

extremal classical probability vectors, corresponding to n_v extremal deterministic LHVMs. Hereafter, we will also refer to the extreme points of a convex polytope \mathcal{P} as its *vertices*, denoted as $\text{vert}(\mathcal{P})$. The fact that there are only a finite number of extreme points in the (convex) set of classical correlations immediately implies that $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ is a *convex polytope* [106, 107], which was first called the *correlation polytope* by Pitowsky [98]. Notice that the dimension of the correlation polytope, i.e., the dimension of its affine hull⁷ is d_p' .

3.2.3 Correlation Polytope and Bell Inequalities

A well-established fact about a convex polytope is that it can equivalently be represented by the intersection of a finite family of closed halfspaces [106, 107]. As is well-known, a closed halfspace in \mathbb{R}^{d_p} can be represented by an inequality that is *linear* in the d_p coordinates. Let

⁶In the context of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ and where measurement outcomes are bounded between 1 and -1 the extreme points correspond to those whereby $E(A_{s_a}, B_{s_b}) = o_a(A_{s_a}) o_b(B_{s_b}) = \pm 1$.

⁷An affine combination of a set of points $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ is a linear combination of x_k , i.e., $\sum_k q_k x_k$ such that $q_k = 1$. The affine hull of \mathcal{X} is the union of all affine combinations of \mathcal{X} .

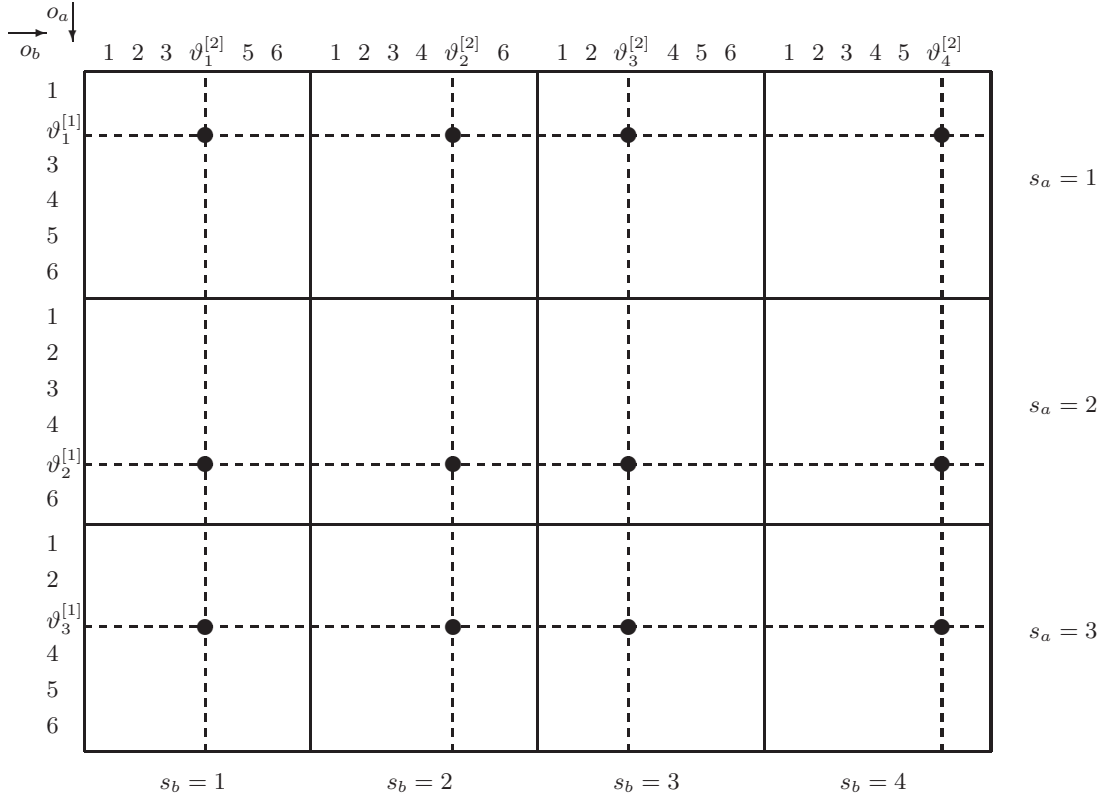


FIGURE 3.1: Schematic representation of the LHV model corresponding to a particular extreme point of $\mathcal{P}_{3;4}^{6;6}$, denoted by $\mathbf{a}, \mathbf{b} \mathbf{B}_{AB}$, where $\mathbf{a} \equiv (\vartheta_1^{[1]}, \vartheta_2^{[1]}, \vartheta_3^{[1]}) = (2, 5, 3)$ and $\mathbf{b} \equiv (\vartheta_1^{[2]}, \vartheta_2^{[2]}, \vartheta_3^{[2]}, \vartheta_4^{[2]}) = (4, 5, 3, 6)$ (adapted from Figure 1 of Ref. [8]). Each row (column), separated from each other by solid horizontal (vertical) lines, corresponds to a choice of measurement s_a (s_b) for Alice (Bob). The intersection of a row and a column gives rise to a sector, which corresponds to particular choice of Alice's and Bob's measurement. For each extremal LHV, the outcome of measurements solely depends on the choice of local measurement. Hence, once a row (column) is chosen, the measurement outcome is also determined, and is indicated by a dashed horizontal (vertical) line. For example, Alice will always observe the second outcome ($o_a = 2$) whenever she chooses to perform the first measurement ($s_a = 1$), regardless of Bob's choice of measurement.

us denote by $I_{\mathbf{m};\mathbf{n}}^{(k)}$ the inequality that is associated with the “ k ”-th halfspace “bounding” the polytope $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, i.e.,

$$I_{\mathbf{m};\mathbf{n}}^{(k)} : \mathbf{F}^{(k)} \cdot \mathbf{p} \leq \beta_{\text{LHV}}^{(k)}, \quad (3.12)$$

then the *boundary* associated with this halfspace is the hyperplane

$$\mathcal{S}_{\text{LHV}}^{(k)}(\mathbf{m}; \mathbf{n}; \mathbf{p}) \equiv \mathbf{F}^{(k)} \cdot \mathbf{p} = \beta_{\text{LHV}}^{(k)}, \quad (3.13)$$

where $\mathbf{p} \in \mathbb{R}^{d_p}$ is an arbitrary vector in the space of correlations, $\mathbf{F}^{(k)}$ is a vector defining the “direction” of the hyperplane involved, $\mathbf{F}^{(k)} \cdot \mathbf{p}$ represents the Euclidean inner product between the two vectors, and $\beta_{\text{LHV}}^{(k)}$ is some constant related to the offset of the hyperplane from the origin.

By definition, a classical probability vector \mathbf{p}_{LHV} is a member of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ and hence must satisfy inequality (3.12), i.e.,

$$\mathbf{F}^{(k)} \cdot \mathbf{p}_{\text{LHV}} \leq \beta_{\text{LHV}}^{(k)}. \quad (3.14)$$

The inequality (3.12) is therefore a valid constraint that has to be satisfied by all classical probability vectors. In other words, it is a *Bell inequality*. It is straightforward to see that any conic combination⁸ of such inequalities will also give rise to another inequality that has to be satisfied by all \mathbf{p}_{LHV} . There is thus no unique family of inequalities defining a given correlation polytope $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$. In principle, one can even write down an infinite family of Bell inequalities, each associated with a closed halfspace (and hence hyperplane), which has to be satisfied by all \mathbf{p}_{LHV} . In this regard, it is worth noting that the smallest family of such closed halfspaces consists of those whose *boundaries* are the affine hull of the facets⁹ of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ (pp 31, [106]). In other words, Bell inequalities that are associated with this smallest family of halfspaces are characterized by \mathbf{F}^k and $\beta_{\text{LHV}}^{(k)}$ such that the solution set $\{\mathbf{p}^{(k)}\} \subset \mathcal{P}_{m_A; m_B}^{n_A; n_B}$ to each of the corresponding equalities

$$\mathbf{F}^{(k)} \cdot \mathbf{p}_{\text{LHV}} = \beta_{\text{LHV}}^{(k)}, \quad (3.15)$$

is nonempty and whose affine dimension equals $d'_p - 1$. For definiteness, we will refer to them as *tight* Bell inequalities [59, 60], or equivalently facet-inducing Bell inequalities [61]. It is worth noting that the coefficients associated with these tight Bell inequalities, i.e., $\mathbf{F}^{(k)}$, when properly normalized, also define a convex polytope that is *dual* to the correlation polytope. Moreover, a probability vector \mathbf{p} is classical if and only if it satisfies this minimal set of Bell inequalities defining $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$.¹⁰

For the convenience of subsequent discussion, let us note that the *linearity* of inequality (3.12) also allows us to write the functional form of a generic Bell inequality, c.f. Eq. (3.13), in the following tensorial form

$$\mathcal{S}_{\text{LHV}}(\mathbf{m}; \mathbf{n}; \mathbf{p}) = \sum_{s_a=0}^{m_A} \sum_{s_b=0}^{m_B} \sum_{o_a=1}^{n_A} \sum_{o_b=1}^{n_B} b_{s_a s_b}^{o_a o_b} p_{s_a s_b}^{o_a o_b} + b_{0,0}, \quad (3.16)$$

where

$$p_{s_a s_b}^{o_a o_b} \equiv \begin{cases} p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b) & : s_a > 0, s_b > 0, \\ p_{\mathcal{A}}^{o_a}(s_a) & : s_a > 0, s_b = 0, \\ p_{\mathcal{B}}^{o_b}(s_b) & : s_a = 0, s_b > 0, \end{cases} \quad (3.17)$$

is a component of the probability vector \mathbf{p} and $b_{s_a s_b}^{o_a o_b}$ is the corresponding component of the vector of coefficients \mathbf{F} . Notice that the sums over outcomes are restricted in that when $s_a = 0$, there is no sum over o_a and when $s_b = 0$, there is no sum over o_b ; in these special cases, we shall write

$$b_{s_a s_b}^{o_a o_b} \equiv \begin{cases} b_{s_a 0}^{o_a} & : s_a > 0, s_b = 0, \\ b_{0 s_b}^{o_b} & : s_a = 0, s_b > 0. \end{cases} \quad (3.18)$$

⁸A conic combination of n points is a non-negative linear combination of the n points.

⁹The intersection of a polytope with a supporting hyperplane gives rise to a face of the polytope. If the dimension of a polytope is d , then a face of dimension $d - 1$ is known as a facet of the polytope.

¹⁰See, for example Definition 2.10 and Theorem 2.15 of Ref. [107].

We can then write these coefficients in a compact manner via the following matrix

$$b \sim \left(\begin{array}{c|c|c|c|c} b_{0,0} & \mathbf{b}_{0,1} & \mathbf{b}_{0,2} & \cdots & \mathbf{b}_{0,m_B} \\ \hline \mathbf{b}_{1,0} & \mathbf{b}_{1,1} & \mathbf{b}_{1,2} & \cdots & \mathbf{b}_{1,m_B} \\ \hline \mathbf{b}_{2,0} & \mathbf{b}_{2,1} & \mathbf{b}_{2,2} & \cdots & \mathbf{b}_{2,m_B} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline \mathbf{b}_{m_A,0} & \mathbf{b}_{m_A,1} & \mathbf{b}_{m_A,2} & \cdots & \mathbf{b}_{m_A,m_B} \end{array} \right), \quad (3.19)$$

where each of the boldfaced entries in the above matrix is a block matrix of appropriate dimension. For example, $\mathbf{b}_{1,1}$ in the above matrix is the following $(n_A - 1) \times (n_B - 1)$ matrix

$$\mathbf{b}_{1,1} \equiv \begin{pmatrix} b_{11}^{11} & b_{11}^{12} & \cdots & b_{11}^{1n_B} \\ b_{11}^{21} & b_{11}^{22} & \cdots & b_{11}^{2n_B} \\ \vdots & \vdots & \vdots & \vdots \\ b_{11}^{n_A-1,1} & b_{11}^{n_A-1,2} & \cdots & b_{11}^{n_A-1,n_B-1} \end{pmatrix}, \quad (3.20)$$

whereas $\mathbf{b}_{1,0}$ and $\mathbf{b}_{0,1}$ are, respectively, column vector and row vector of length $n_A - 1$ and $n_B - 1$.¹¹ It is then expedient to write a Bell inequality explicitly as

$$I_{\mathbf{m};\mathbf{n}}^{(k)} : \mathcal{S}_{\text{LHV}}^{(k)}(\mathbf{m}; \mathbf{n}; \mathbf{p}) \leq \beta_{\text{LHV}}^{(k)}, \quad (3.21)$$

but compactly as

$$I_{\mathbf{m};\mathbf{n}}^{(k)} : b^{(k)} \leq \beta_{\text{LHV}}^{(k)}, \quad (3.22)$$

where $b^{(k)}$ is the corresponding matrix of coefficients, Eq. (3.19) – Eq. (3.20), associated with the specific Bell inequality.

As an example, let us look at the simplest nontrivial scenario where $\mathbf{m} = (2, 2)$ and $\mathbf{n} = (2, 2)$. In this case, it is known for a long time [60, 62, 99, 109] that the only class of nontrivial¹² tight Bell inequalities are the Bell-CH inequalities listed in (2.32a) and their equivalents. In this case, we have¹³

$$\mathcal{S}_{\text{LHV}}^{(\text{CH})} = p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{A}}^{o_a}(1) - p_{\mathcal{B}}^{o_b}(1) \leq 0. \quad (3.23)$$

Making use of the matrix representation introduced above, we will write this class of inequalities as [60]

$$I_{(2,2);(2,2)}^{(\text{CH})} : \left(\begin{array}{c|c|c} \cdot & -1 & \cdot \\ \hline -1 & 1 & 1 \\ \hline \cdot & 1 & -1 \end{array} \right) \leq 0, \quad (3.24)$$

where for ease of reading, we will always replace each null entry in a matrix by a single dot.

¹¹The length of these vectors as well as the dimension of each block matrix can be traced back to the discussion around Eq. (3.5). Consequently, for a two-outcome Bell inequality (e.g. the Bell-CH inequality), or a Bell correlation inequality, we will collapse each block matrix and write it as a single number.

¹²The other tight Bell inequalities are trivial in the sense that they either require probabilities to be non-negative or not larger than unity.

¹³Hereafter, we will drop the arguments of \mathcal{S}_{LHV} for brevity of notation.

A great advantage of this matrix representation is that a Bell inequality that only differs from another in its label of *measurement settings* can be obtained from (the matrix representation of) the other by applying an appropriate permutation to the rows and/or columns of blocks in the associated matrix of coefficients, c.f. Eq. (3.19). Similarly, two Bell inequalities that differ from another only in their label of *measurement outcomes* for a particular local measurement setting can be obtained from one another by applying appropriate permutation to the rows and/or columns within the entire row/column of *blocks of matrix of coefficients*, c.f. Eq. (3.20). And finally, two Bell inequalities that only differ in their label of observers, e.g. “Alice” \leftrightarrow “Bob”, can be obtained from one another by transposing their respective matrix of coefficients (see Appendix B.1.1 for examples). With this compact notation, the stage is now set for us to look into Bell inequalities that arise in the more complicated experimental scenarios.

3.3 The Zoo of Bell Inequalities

To date, a zoo of Bell inequalities is available in the literature. In particular, a handful of these were constructed in the 1980s [97, 102, 110, 111, 112, 113, 114] primarily to investigate if Bell inequality violation would vanish in one of the plausible classical limits. Some of these early attempts, however, suffered by their rather *ad hoc* construction of (non-tight) Bell inequalities. In what follows, we will review, via the characterization of various classical correlation polytopes $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, some of the more well-known (tight) Bell inequalities beyond Bell-CH and Bell-CHSH that can be, or have been constructed using the geometrical approach presented above.

For bipartite Bell inequalities, that is, Bell inequalities involving only two parties, our discussion will be carried out primarily for Bell inequalities for probabilities, as this is where most of the work was done [14, 60, 109]. On the contrary, most of the work for multipartite Bell inequalities were carried out in the context of Bell correlation inequalities, in particular those involving only the full correlation functions.¹⁴

3.3.1 Other Bipartite Bell Inequalities for Probabilities

3.3.1.1 Two Outcomes $\mathbf{n} = (2, 2)$

Now, let us focus on bipartite Bell inequalities for probabilities involving only dichotomic observables, i.e., $\mathbf{n} = (2, 2)$. For scenarios involving more than two measurements on one side, but not on the other, i.e., $\mathbf{m} = (2, m)$ or $\mathbf{m} = (m, 2)$ with $m > 2$, Collins and Gisin have shown in Ref. [60] that there are no new tight Bell inequalities. In other words, all facets of the correlation polytope $\mathcal{P}_{2;m}^{2;2}$ (equivalently $\mathcal{P}_{m;2}^{2;2}$) either correspond to the trivial requirement of probabilities being positive, or to a Bell-CH type inequality involving only two out of the m possible measurements. An example of such an inequality would be [60]

$$p_{AB}^{o_a o_b}(1, 1) + p_{AB}^{o_a o_b}(1, m) + p_{AB}^{o_a o_b}(2, 1) - p_{AB}^{o_a o_b}(2, m) - p_A^{o_a}(1) - p_B^{o_b}(1) \leq 0, \quad (3.25)$$

¹⁴A full correlation function, as opposed to a restricted correlation function, for an N -party Bell experiment is a correlation function that takes the local observables at all the N sites as arguments [115] (see the discussion at pp. 30 for more details).

in which case only statistics of Bob's first and m^{th} local measurement are involved in the above inequality.

In the case when each party is allowed to perform three alternative measurements, i.e., for the correlation polytope $\mathcal{P}_{3;3}^{2;2}$, a complete list of 684 facets was first obtained by Pitowsky and Svozil in Ref. [109]. Among these, Collins and Gisin [60] have found that there are 36 positive probability facets, 72 Bell-CH-type facets while the remaining facets are associated with inequalities that are equivalent to¹⁵

$$I_{(3,3);(2,2)}^{(1)} : \left(\begin{array}{c|c|c|c|c} \cdot & -2 & -1 & \cdot & \\ \hline -1 & 1 & 1 & 1 & \\ \hline \cdot & 1 & 1 & -1 & \\ \hline \cdot & 1 & -1 & \cdot & \end{array} \right) \leq 0. \quad (3.26)$$

Equivalently, in the notation of Eq. (3.17), inequality $I_{(3,3);(2,2)}^{(1)}$ can be written more explicitly as:

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_{3322})} = & p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(1, 3) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) + p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{AB}}^{o_a o_b}(2, 3) \\ & + p_{\mathcal{AB}}^{o_a o_b}(3, 1) - p_{\mathcal{AB}}^{o_a o_b}(3, 2) - p_{\mathcal{A}}^{o_a}(1) - 2p_{\mathcal{B}}^{o_b}(1) - p_{\mathcal{B}}^{o_b}(2) \leq 0, \end{aligned} \quad (3.27)$$

which is understood to hold true for arbitrary but fixed choice of o_a and o_b .

Here, we again see that a lower dimensional Bell inequality, namely, the Bell-CH inequality occurring as a facet of a more complicated correlation polytope. As was shown by Pironio [117], this is actually a generic feature of tight Bell inequalities for probabilities¹⁶, i.e., when lifted to a more complicated experimental scenario, say, involving more local measurement settings and/or outcomes and/or number of parties, the lower dimensional Bell inequality will still serve as a tight Bell inequality in the higher dimensional space. Since a direct enumeration of all tight Bell inequalities is computationally intensive and may not be feasible in practice, this property of tight Bell inequalities will enable us to find out, at least, a partial list of facets in the higher dimensional correlation polytope [117].

For example, for the correlation polytope $\mathcal{P}_{4;4}^{2;2}$, even though a complete characterization of tight Bell inequalities for probabilities is not known, we do know from Pironio's result [117] that all the tight inequalities derived from, say, $\mathcal{P}_{3;4}^{2;2}$ will also serve as tight inequalities in the higher-dimensional space. This lower dimensional case has been fully characterized in Ref. [60] and the correlation polytope $\mathcal{P}_{3;4}^{2;2}$ is known to be made up of from five different classes of facets.

For $\mathcal{P}_{4;4}^{2;2}$, however, it is known that there are also other classes of tight Bell inequalities. For example, Collins and Gisin [60] have shown that a generalization of $I_{(3,3);(2,2)}^{(1)}$, namely,

$$I_{(4,4);(2,2)}^{(3)} : \left(\begin{array}{c|c|c|c|c} \cdot & -3 & -2 & -1 & \cdot \\ \hline -1 & 1 & 1 & 1 & 1 \\ \hline \cdot & 1 & 1 & 1 & -1 \\ \hline \cdot & 1 & 1 & -1 & \cdot \\ \hline \cdot & 1 & -1 & \cdot & \cdot \end{array} \right) \leq 0, \quad (3.28)$$

¹⁵The analogous analysis for Bell correlation inequalities with $\mathbf{m} = (3, 3)$ has also been carried out independently by Śliwa [116] (see also [96]).

¹⁶See Avis *et al.* [58] for the analogous proof for Bell correlation inequalities.

is also a tight Bell inequality. By brute force, Ito *et al.* [14] have found two other tight Bell inequalities for this experimental scenario:

$$I_{(4,4);(2,2)}^{(1)} : \left(\begin{array}{c|c|c|c|c} \cdot & \cdot & -1 & -1 & -1 \\ \hline -1 & -1 & 1 & \cdot & 2 \\ \hline \cdot & \cdot & 1 & -1 & -1 \\ \hline -1 & 1 & -1 & 1 & 1 \\ \hline -1 & -1 & 1 & 2 & -1 \end{array} \right) \leq 0, \quad I_{(4,4);(2,2)}^{(2)} : \left(\begin{array}{c|c|c|c|c} \cdot & \cdot & \cdot & -1 & -1 \\ \hline \cdot & 1 & 1 & 1 & \cdot \\ \hline -1 & 1 & -1 & \cdot & 1 \\ \hline -1 & -1 & 1 & 1 & 1 \\ \hline \cdot & \cdot & -1 & 1 & \cdot \end{array} \right) \leq 0, \quad (3.29)$$

and by the method of *triangular elimination* [61], they have also found at least one other tight Bell inequality:

$$I_{(4,4);(2,2)}^{(4)} : \left(\begin{array}{c|c|c|c|c} \cdot & \cdot & -1 & -1 & -1 \\ \hline -2 & -1 & 1 & \cdot & 2 \\ \hline -1 & \cdot & 1 & -1 & -1 \\ \hline -1 & 1 & -1 & 1 & 1 \\ \hline \cdot & -1 & 1 & 2 & -1 \end{array} \right) \leq 0. \quad (3.30)$$

which they have labeled as “A5”. Very recently, a partial list of 26 inequivalent facet-inducing inequalities for $\mathcal{P}_{4,4}^{2,2}$ was presented in Ref. [118].

Beyond this, a systematic characterization of all the tight Bell inequalities with more local measurements seems formidable. However, we do know that both $I_{(3,3);(2,2)}^{(1)}$ and $I_{(4,4);(2,2)}^{(3)}$ are members of a broader class of Bell inequalities, called I_{mm22} by Collins and Gisin [60]. It is worth noting that this class of inequalities is *asymmetric* with respect to swapping Alice and Bob. In particular, for $\mathbf{m} = (m, m)$, the inequality admits the following compact representation [60, 61]:¹⁷

$$I_{mm22} : \left(\begin{array}{c|c|c|c|c|c|c|c|c} \cdot & -1 & \cdot & \cdot & \dots & \cdot & \cdot & \cdot & \cdot \\ \hline -(m-1) & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1 \\ \hline -(m-2) & 1 & 1 & 1 & \dots & 1 & 1 & 1 & -1 \\ \hline -(m-3) & 1 & 1 & 1 & \dots & 1 & 1 & -1 & \cdot \\ \hline -(m-4) & 1 & 1 & 1 & \dots & 1 & -1 & \cdot & \cdot \\ \hline \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \hline \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \hline -1 & 1 & 1 & -1 & \cdot & \dots & \cdot & \cdot & \cdot \\ \hline \cdot & 1 & -1 & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \end{array} \right) \leq 0. \quad (3.31)$$

For $m \leq 7$, Collins and Gisin computationally verified that each I_{mm22} is a tight Bell inequality, and for general m , the tightness of these inequalities has also been proven very recently by Avis and Ito [119]. Apart from this, Avis *et al.* [61] have also obtained a huge

¹⁷Note that for the specific case of $m = 3$ and $m = 4$, Eq. (3.31) is related to, respectively, Eq. (3.26) and Eq. (3.28) by a transposition, which corresponds to swapping the label “Alice” \leftrightarrow “Bob”. The current form of Eq. (3.31), as opposed to Eq. (3.26) and Eq. (3.28), looks closer to the original form presented in Refs. [60, 61].

number of tight Bell inequalities by applying the method of triangular elimination to a list of tight inequalities for the so-called *cut polytope*. On top of inequality (3.30), the explicit form of some of these inequalities with $m_A, m_B \leq 5$ can also be found in Ref. [14, 120].

3.3.1.2 More than Two Outcomes

The set of classical correlations involving greater number of measurement outcomes is apparently not as well known. In particular, investigation carried out by Collins and Gisin [60] suggests that for $2 < n \leq 5$, all facets of the correlation polytope $\mathcal{P}_{2;2}^{2;n}$ (equivalently $\mathcal{P}_{2;2}^{n;2}$) are either of the Bell-CH-type or the trivial type that requires non-negativity of probabilities. For $\mathcal{P}_{2;2}^{3;3}$, it was shown by Masanes [59] that there is only one other class of tight Bell inequalities, which was discovered independently by Collins *et al.* [121] and Kaslikowski *et al.* [122]. Following Ref. [60], we will write this inequality as

$$I_{(2,2);(3,3)}^{(1)} : \left(\begin{array}{c|c|c} \cdot & -\mathbf{1}_2^T & \mathbf{0}_2^T \\ \hline -\mathbf{1}_2 & M_1 & M_2 \\ \hline \mathbf{0}_2 & M_2 & -M_2 \end{array} \right) \leq 0, \quad (3.32)$$

where $\mathbf{1}_2$ and $\mathbf{0}_2$ are, respectively, column vector of ones and zeros with length 2,

$$M_1 \equiv \begin{pmatrix} 1 & 1 \\ 1 & \cdot \end{pmatrix}, \quad M_2 \equiv \begin{pmatrix} \cdot & 1 \\ 1 & 1 \end{pmatrix}. \quad (3.33)$$

In Ref. [60], the inequality (3.32) was actually presented as a special case of a class of inequalities — which Collins and Gisin labeled as I_{22nn} — that holds for arbitrary $\mathbf{n} = (n, n)$. Specifically, for $n = 4$, it takes the form of

$$I_{(2,2);(4,4)}^{(1)} : \left(\begin{array}{c|c|c} & -1 & -1 & -1 & \cdot & \cdot & \cdot \\ \hline -1 & 1 & 1 & 1 & \cdot & \cdot & 1 \\ -1 & 1 & 1 & \cdot & \cdot & 1 & 1 \\ -1 & 1 & \cdot & \cdot & 1 & 1 & 1 \\ \hline \cdot & \cdot & \cdot & 1 & \cdot & \cdot & -1 \\ \cdot & \cdot & 1 & 1 & \cdot & -1 & -1 \\ \cdot & 1 & 1 & 1 & -1 & -1 & -1 \end{array} \right) \leq 0, \quad (3.34)$$

where inequalities for higher values of n involve the obvious modifications on individual blocks. For general n , we can write I_{22nn} in the following functional form:

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_{22nn})} &= \sum_{o_a=1}^{n-1} \sum_{o_b=1}^{n-o_a} p_{AB}^{o_a o_b}(1, 1) + \sum_{o_a=1}^{n-1} \sum_{o_b=n-o_a}^{n-1} \left[p_{AB}^{o_a o_b}(1, 2) + p_{AB}^{o_a o_b}(2, 1) - p_{AB}^{o_a o_b}(2, 2) \right] \\ &\quad - \sum_{o_a=1}^{n-1} p_A^{o_a}(1) - \sum_{o_b=1}^{n-1} p_B^{o_b}(1) \leq 0. \end{aligned} \quad (3.35)$$

This class of inequalities is believed [60] to be equivalent to the more well-known n -outcome Collins-Gisin-Linden-Massar-Popescu (henceforth abbreviated as CGLMP) inequality [121], which admits the following functional form:¹⁸

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_n)} = \sum_{k=0}^{\lfloor \frac{n}{2}-1 \rfloor} \left(1 - \frac{2k}{n-1} \right) \sum_{o_b=1}^n & \left[p_{\mathcal{AB}}^{o_b-k o_b}(1, 1) - p_{\mathcal{AB}}^{o_b+k+1 o_b}(1, 1) + p_{\mathcal{AB}}^{o_b+k o_b}(1, 2) - p_{\mathcal{AB}}^{o_b-k-1 o_b}(1, 2) \right. \\ & \left. + p_{\mathcal{AB}}^{o_b+k o_b}(2, 1) - p_{\mathcal{AB}}^{o_b-k-1 o_b}(2, 1) + p_{\mathcal{AB}}^{o_b-k-1 o_b}(2, 2) - p_{\mathcal{AB}}^{o_b+k o_b}(2, 2) \right] \leq 2, \end{aligned} \quad (3.36)$$

where expression such as $o_b - k$ in the above inequality is understood to be evaluated modulo n . A proof of their equivalence is, however, not available in the literature. In Appendix B.1.1, we have provided this missing proof. The tightness of the CGLMP inequality, and hence I_{22nn} was proven by Masanes in Ref. [59]. They therefore correspond to facets of $\mathcal{P}_{2;2}^{n;n}$ for arbitrary $n \geq 2$.

Finally, we note that a family of (tight) Bell inequalities – the I_{mmnn} inequality — involving more than two measurements per site, and more than two outcomes per measurement has also been presented in Ref. [60]. However, the I_{mmnn} inequality is only known to correspond to facets of $\mathcal{P}_{m;m}^{n;n}$ with $m, n > 2$ for some relatively small values of m and n .

3.3.2 Other Bipartite Correlation Inequalities

Thus far, we have focused on the analysis of correlation polytopes living in the space of probability vectors $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$ and looked at the corresponding tight Bell inequalities bounding these polytopes. Now, let us turn our attention to the space of correlations defined in terms of correlation functions – denoted by ${}^c\mathcal{C}_{m_A; m_B}^{2;2}$ — for an experimental scenario involving only two parties performing m_A and m_B dichotomic¹⁹ measurements, and whose measurement outcomes are labeled by ± 1 . Specifically, in this bipartite scenario, ${}^c\mathcal{C}_{m_A; m_B}^{2;2}$ is a space of dimension $d_c = m_A m_B + m_A + m_B$, which can be labeled by the following coordinates [58, 96]

$$\begin{aligned} \{ & E(A_1, B_1), \dots, E(A_1, B_{m_B}), E(A_2, B_1), \dots, E(A_{m_A}, B_{m_B}), \\ & E(A_1), \dots, E(A_{m_A}), E(B_1), \dots, E(B_{m_B}) \}, \end{aligned} \quad (3.37)$$

where the *restricted* correlation functions [115] are defined as

$$E(A_{s_a}) = p_{\mathcal{A}}^+(s_a) - p_{\mathcal{A}}^-(s_a), \quad E(B_{s_b}) = p_{\mathcal{B}}^+(s_b) - p_{\mathcal{B}}^-(s_b). \quad (3.38)$$

More often than not, however, we are only interested in the (sub)space of correlations that is defined solely in terms of the *full* correlation functions $E(A_{s_a}, B_{s_b})$. We shall denote this subspace by ${}^s\mathcal{C}_{m_A; m_B}^{2;2}$. Note that it is a subspace of dimension $d_s = m_A m_B$. As with $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$, the set of classical correlations in ${}^c\mathcal{C}_{m_A; m_B}^{2;2}$ (${}^s\mathcal{C}_{m_A; m_B}^{2;2}$) is a convex polytope which we shall denote by ${}^c\mathcal{P}_{m_A; m_B}^{2;2}$ (${}^s\mathcal{P}_{m_A; m_B}^{2;2}$).

¹⁸Here, we have swapped B_1 and B_2 (i.e., Bob's first and second measurement settings) and followed Ref. [59] by grouping terms for the same setting together. Moreover, we have also shifted the constant “2” to the *lhs* of the inequality.

¹⁹Strictly, many of the subsequent discussion will still hold true even if we have more outcomes in the experiments, provided that all measurement outcomes are bounded between “−1” and “1”.

A well-known fact in relation to these polytopes is that the two polytopes ${}^c\mathcal{P}_{m_A; m_B}^{2;2}$ and $\mathcal{P}_{m_A; n_A}^{2;2}$ are actually isomorphic (see for example [58]). Therefore, any tight Bell inequality defining $\mathcal{P}_{m_A; n_A}^{2;2}$ can also be mapped to a tight correlation inequality defining ${}^c\mathcal{P}_{m_A; m_B}^{2;2}$ via Eq. (2.34) and Eq. (3.38). Nevertheless, for the purpose of performing this mapping, it is more convenient to make use of an equivalent form of Eq. (2.34),

$$E(A_{s_a}, B_{s_b}) = 1 - 2p_A^+(s_a) - 2p_B^+(s_b) + 4p_{AB}^{++}(s_a, s_b). \quad (3.39)$$

For instance, in the simplest scenario of $\mathbf{m} = (2, 2)$, one obtains the Bell-CHSH inequality from the Bell-CH inequality via Eq. (3.39). Similarly, by applying Eq. (3.38) and Eq. (3.39) to $I_{(3,3);(2,2)}^{(1)}$, Eq. (3.27), one can obtain the following correlation inequality [58, 116]

$$\begin{aligned} &E(A_1, B_1) + E(A_1, B_2) + E(A_1, B_3) + E(A_2, B_1) + E(A_2, B_2) - E(A_2, B_3) \\ &+ E(A_3, B_1) - E(A_3, B_2) - E(A_1) - E(A_2) + E(B_1) + E(B_2) \leq 4. \end{aligned} \quad (3.40)$$

Note, nonetheless, that as opposed to the Bell-CHSH inequality, inequality (3.40) does not live in the subspace of full correlations ${}^s\mathcal{C}_{3;3}^{2;2}$, i.e., it *is not* a facet-inducing inequality for ${}^s\mathcal{P}_{3;3}^{2;2}$. In fact, recent work by Avis *et al.* [58] has demonstrated that for $\mathbf{m} = (m_A, m_B)$ with $\min\{m_A, m_B\} \leq 3$, the Bell-CHSH inequalities and the trivial inequalities²⁰

$$-1 \leq E(A_{s_a}, B_{s_b}) \leq 1, \quad (3.41)$$

for all $s_a = 1, \dots, m_A$ and all $s_b = 1, \dots, m_B$, are the only tight correlation inequalities defining ${}^s\mathcal{P}_{3;3}^{2;2}$.

On the contrary, when four alternative measurements are allowed at each site, Gisin has constructed the following Bell correlation inequalities [53]

$$AS_4 : \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -2 & \cdot \\ 1 & -1 & \cdot & \cdot \end{pmatrix} \leq 6, \quad (3.42)$$

$$D_4 : \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 1 & 2 & -2 \\ 1 & 2 & -2 & -1 \\ 2 & -2 & -1 & -1 \end{pmatrix} \leq 10, \quad (3.43)$$

where the (s_a, s_b) entry in each matrix is the coefficient associated with the full correlation function $E(A_{s_a}, B_{s_b})$. These inequalities are tight. Together with the trivial inequalities (3.41) and the Bell-CHSH inequality, they form a complete set of tight correlation inequalities defining ${}^s\mathcal{P}_{4;4}^{2;2}$ [58].

As a last remark, we note that the correlation inequality AS_4 has been generalized to an arbitrary even number of measurement settings. Moreover, they can also be seen as a correlation inequality that is valid for arbitrary number of measurement outcomes if instead

²⁰In relation to Eq. (3.39) and Eq. (3.38), these trivial inequalities are in one-to-one correspondence with the trivial requirement of probabilities being non-negative and less than or equal to one.

of Eq. (2.34) or Eq. (3.39), which are only for two-outcome Bell experiments, the full correlation function $E(A_{s_a}, B_{s_b})$, Eq. (2.24), is interpreted as the difference between the probability of observing the same outcomes at the two sites and the probability of observing different outcomes at the two sites, as it was done in Eq. (2.35) [53].

3.3.3 Multipartite Bell Inequalities

In sharp contrast with the study of bipartite Bell inequalities — where most developments were carried out in the context of *probability vectors* — the multipartite analog was primarily developed in the context of *correlation functions*, and in particular the full correlation functions. The pioneering work in this regard was initiated by Mermin [123] who, in turn, was inspired by the results presented by Greenberger, Horne and Zeilinger (henceforth abbreviated as GHZ) on a demonstration of incompatibility between local causality and quantum mechanical prediction without resorting to any inequalities [124, 125].

In his seminal work, Mermin [123] investigated a scenario involving n parties and where each of them can perform two dichotomic measurements. Starting from the assumption of a general LHV, Mermin constructed a Bell correlation inequality which involves only n -partite full correlation functions; his inequality therefore defines a closed halfspace in ${}^s\mathcal{C}_{2;2;\dots;2}^{2;2;\dots;2}$ where ${}^s\mathcal{P}_{2;2;\dots;2}^{2;2;\dots;2}$ resides (here, there are n indices in both the superscript²¹ and subscript). This work was further developed by Roy and Singh [126], Ardehali [127], and eventually by Belinskiĭ and Klyskho [128, 129] whereby the current form of Mermin inequality (also commonly known as Mermin-Ardehali-Belinskiĭ-Klyskho, or in short, MABK inequality) was culminated.²²

An interesting feature of the present form of Mermin's inequality is that all inequalities involving $n > 2$ parties can be obtained from the Bell-CHSH inequality in a recursive manner. To see that, let us now denote by $o_{s_j}^{[j]} = \pm 1$ the outcome of measurement when the j^{th} observer chooses to measure the s_j^{th} dichotomic observable. As a classical variable,²³ $o_{s_j}^{[j]}$ can be defined independently for each j and each s_j . Thus, in each run of the experiment, the expression

$$F_2 \equiv \frac{1}{2} \left(o_1^{[1]} + o_2^{[1]} \right) o_1^{[2]} + \frac{1}{2} \left(o_1^{[1]} - o_2^{[1]} \right) o_1^{[2]}, \quad (3.44)$$

must either end up as 1 or -1 , since either $o_1^{[1]} = o_2^{[1]}$ or $o_1^{[1]} = -o_2^{[1]}$. Averaging this expression over many runs of the experiment, we see that the average value of F_2 must be less than or equal to 1, since each term in the average is at most 1. This is essentially a statement of the Bell-CHSH inequality given in Eq. (2.26).

To obtain the n -partite Mermin's inequality, we now follow Ref. [130] and define

$$F_n \equiv \frac{1}{2} \left(o_1^{[n]} + o_2^{[n]} \right) F_{n-1} + \frac{1}{2} \left(o_1^{[n]} - o_2^{[n]} \right) F'_{n-1}, \quad (3.45)$$

²¹We are generalizing the notation introduced in Sec. 3.3.1 such that indices in the subscript (sequentially) indicate the number of possible measurements at each site and indices in the superscript indicate the number of possible outcomes per measurement at each site.

²²The inequality developed by Roy and Singh [126] is actually equivalent to the current form of Mermin's inequality developed by Belinskiĭ and Klyskho [128, 129].

²³That is, a variable that can be defined using local hidden-variable.

where F'_{n-1} is the same expression as F_{n-1} except that all the $o_1^{[j]}$ and $o_2^{[j]}$ are interchanged. By going through the same reasoning as before, it is not difficult to see that the average value of F_n must be bounded above by 1, i.e.,

$$\text{Exp} \left(F_n(o_1^{[1]}, \dots, o_1^{[n]}, o_2^{[1]}, \dots, o_2^{[n]}) \right) \leq 1, \quad (3.46)$$

where here, $\text{Exp}(x)$ refers to the expectation value of x . It is also not difficult to see from Eq. (3.45) that F_n is an expression that is linear in all the local variable $o_{s_j}^{[j]}$, therefore by generalizing the notation introduced in Eq. (3.16), we can write

$$F_n = \sum_{s_1, s_2, \dots, s_n=1}^2 b_{s_1 s_2 \dots s_n} \prod_{j=1}^n o_{s_j}^{[j]}, \quad (3.47)$$

for some specific $b_{s_1 s_2 \dots s_n}$. Now, we can write the entire class of Mermin inequalities in a form that is closer to inequality (2.26), i.e.,

$$\sum_{s_1=1}^2 \cdots \sum_{s_n=1}^2 b_{s_1 s_2 \dots s_n} E(o_{s_1}^{[1]}, o_{s_2}^{[2]}, \dots, o_{s_n}^{[n]}) \leq 1, \quad (3.48)$$

where $E(\cdot)$ is the n -party correlation function defined analogous to Eq. (2.24).

As is now well-known, Mermin inequalities are not the only class of n -partite Bell correlation inequalities. In fact, a *complete*²⁴ set of 2^{2^n} Bell correlation inequalities involving only the full correlation functions has been obtained independently by Werner & Wolf [115] and Żukowski & Brukner [131] (the complete set of inequalities for $n = 4$ was also obtained by Weinfurter and Żukowski in Ref. [132]). All these inequalities are uniquely characterized by the tensor $b_{s_1 s_2 \dots s_n}$, which can be written as [115]

$$b_{s_1 s_2 \dots s_n} = 2^{-n} \sum_{r_1=0}^1 \sum_{r_2=0}^1 \cdots \sum_{r_n=0}^1 f(r_1, r_2, \dots, r_n) (-1)^{\sum_j r_j (s_j - 1)} \quad (3.49)$$

where $f(r_1, r_2, \dots, r_n) \in \{+1, -1\}$ is a binary function that takes an n -bit-vector r (with components r_i) as argument. There are altogether 2^{2^n} of such functions, each of them gives rise to a unique tensor $b_{s_1 s_2 \dots s_n}$ which, in turn, defines a Bell correlation inequality via Eq. (3.48). These inequalities are tight [115, 131], and therefore are facet inducing for the correlation polytope of n -partite correlation functions ${}^s\mathcal{P}_{2;2;\dots;2}^{2;2;\dots;2}$. It happens that ${}^s\mathcal{P}_{2;2;\dots;2}^{2;2;\dots;2}$ is actually a 2^n -dimensional hyperoctahedron [115], and hence the complete set of 2^{2^n} inequalities is equivalent to a single nonlinear inequality [115, 131].

More recently, by generalizing the work of Wu and Zong on ${}^s\mathcal{P}_{4;2;2;\dots;2}^{2;2;2;\dots;2}$ [133], Laskowski and coworkers [134] have come up with a systematic way to generate a huge class of tight Bell correlation inequalities for ${}^s\mathcal{P}_{2^{n-1};2^{n-1};2^{n-2};2^{n-3};\dots;2}^{2;2;2;\dots;2}$. In particular, explicit forms of these facet-inducing inequalities for ${}^s\mathcal{P}_{4;4;2}^{2;2;2}$ and ${}^s\mathcal{P}_{8;8;4;2}^{2;2;2;2}$ can be found in Ref. [134]. A first step towards

²⁴Complete, in the sense that a vector of full correlation functions is classical if and only if it satisfies all of these inequalities.

the complete characterization of facets for a more symmetrical experimental scenario, namely, ${}^s\mathcal{P}_{3;3;3;\dots;3}^{2;2;2;\dots;2}$ was carried out in Ref. [135] by Żukowski. Apparently, a complete characterization for this experimental scenario has subsequently been achieved in Ref. [136]. Based on these findings, the explicit form of a tight correlation inequality for ${}^s\mathcal{P}_{3;3;3}^{2;2;2}$ has very recently been derived and presented in Ref. [137].

Finally, we note that at present, only one facet-inducing inequality for $\mathcal{P}_{2;2;2}^{3;3;3}$ is known, and is presented in the form of a coincidence Bell inequality [138]. Other multipartite Bell inequalities, such as those involving restricted correlation functions [139, 140, 141] or in the form of probability inequality [37] can also be found in the literature. Their tightness, however, is not well studied.

3.4 Conclusion

In this chapter, we have looked at the set of classical correlations, i.e., correlations (either in the form of probability vector or a vector of correlation functions) that are describable within the framework of LHVTs and how it is related to the zoo of Bell inequalities that one can find in the literature. Equipped with a solid understanding of the set of classical correlations, we will next investigate what quantum mechanics has to offer, both in terms of classical correlations and correlations that cannot be accounted for using any locally causal theory.

4

Quantum Correlations and Locally Causal Quantum States

In the last chapter, we have looked at the set of classical correlations and the characterization of its boundaries in terms of Bell inequalities. Now, in this chapter, we will move on to study the set of quantum correlations and see how they are related to the set of classical correlations. Some well-known examples of quantum states admitting locally causal description will also be reviewed.

4.1 Introduction

In a nutshell, *quantum correlations* are simply points in the space of correlations, c.f. Sec. 3.2.1, that are realizable by quantum mechanics through some choice of quantum states and some local measurement operators. Ironically, despite the statistical nature of quantum predictions, there was no known study on this specific aspect of quantum predictions prior to the seminal work by Bell in 1964 [5].

After that, it seems to have taken another 16 years before the first quantitative study on the set of quantum correlations was carried out by Tsirelson¹ [143]. In his work [143], Tsirelson showed that the set of quantum correlations in ${}^s\mathcal{C}_{2;2}^{2;2}$ is also bounded by some very similar linear inequalities like its classical partner. However, these linear inequalities (often known as the Tsirelson inequalities) are, in general, not sufficient to distinguish a correlation that is realizable by quantum mechanics from one that is not. In fact, it took a few more years before Tsirelson came up with a set of necessary and sufficient conditions — in terms of inequalities that are non-linear in the correlation functions — for the realizability of a point in ${}^s\mathcal{C}_{2;2}^{2;2}$ using quantum mechanics [144, 145].

¹Incidentally, in response to a question raised by A. M. Vershik (see pp. 884 of Ref. [142]).

Meanwhile, a general study on the structure of the set of quantum correlations beyond the simplest scenario of $m_A = m_B = n_A = n_B = 2$ was taken up by Pitowsky [146]. In fact, it was in Ref. [146] that the convexity of this set and its relationship with the set of classical correlations were, for the first time, formally established (see also Ref. [62]).

From Bell's theorem [5], we have learned that there are quantum correlations that fall outside the classical correlation polytope. A characterization of quantum states that can give rise to such nonclassical correlations is, nevertheless, still lacking. The seminal work by Werner [29] has established that *entanglement* between spatially separated subsystems is a necessary condition to establish nonclassical correlation. Nonetheless, in the same article [29], Werner also provided an example of an entangled state which does not violate any Bell inequalities if the source is directly subjected to local, projective measurements without any preprocessing. In fact, there are now a few known examples of entangled quantum states which admit an explicit LHV [29, 31, 32, 33].

In this chapter, we will start off, in Sec. 4.2, by reviewing some well-known facts about the set of quantum correlations. In the same section, we will also specify what we mean by a *standard Bell experiment*, a key notion that is used in this, as well as the subsequent chapters. After that, in Sec. 4.3, we will review some of the well-known examples of quantum states admitting either a partial, or a full LHV for projective or generalized measurements given by positive-operator-valued measures (POVM).

4.2 Quantum Correlations

Consider again the set of two-party correlations that respects the relativistic causality condition, $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$. In analogy with the idea of a classical probability vector introduced in Chapter 3, we will now define a *quantum probability vector* as follows.²

Definition 5. A probability vector \mathbf{p}_{QM} in $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$ is said to be a *quantum probability vector* if there exists a bipartite quantum state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, i.e., $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, and some (local) POVM elements $A_{s_a}^{o_a} \in \mathcal{B}(\mathcal{H}_A)$, $B_{s_b}^{o_b} \in \mathcal{B}(\mathcal{H}_B)$, i.e., operators satisfying

$$\sum_{o_a=1}^{n_A} A_{s_a}^{o_a} = \mathbf{1}_{d_A} \quad \text{and} \quad \sum_{o_b=1}^{n_B} B_{s_b}^{o_b} = \mathbf{1}_{d_B} \quad \forall \quad s_a, s_b, \quad (4.1a)$$

$$A_{s_a}^{o_a} \geq 0, \quad B_{s_b}^{o_b} \geq 0 \quad \forall \quad s_a, s_b, o_a, o_b, \quad (4.1b)$$

such that the components of the probability vector satisfy

$$p_{AB}^{o_a o_b}(s_a, s_b) = \text{tr}(\rho A_{s_a}^{o_a} \otimes B_{s_b}^{o_b}) \quad (4.2a)$$

$$p_A^{o_a}(s_a) = \text{tr}(\rho A_{s_a}^{o_a} \otimes \mathbf{1}_{d_B}), \quad p_B^{o_b}(s_b) = \text{tr}(\rho \mathbf{1}_{d_A} \otimes B_{s_b}^{o_b}), \quad (4.2b)$$

where $d_A = \dim(\mathcal{H}_A)$ and $d_B = \dim(\mathcal{H}_B)$.

²The definition is given for probability vectors considered in a bipartite correlation experiment and where correlations are expressed in terms of probability vectors. Nonetheless, it should be clear as to how this definition can be generalized to the multipartite scenario, or the space of correlations defined in terms of correlation functions.

Note that in the above definition of a quantum probability vector \mathbf{p}_{QM} , the dimension of the Hilbert spaces is not fixed *a priori*. In other words, the dimension of the Hilbert spaces involved may vary depending on the given probability vector. As with a classical probability vector, we shall also refer to a quantum probability vector, loosely, as a *quantum correlation*. Moreover, the set of quantum correlations will be denoted by $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$.

Physically, a quantum probability vector $\mathbf{p}_{\text{QM}} \in \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ is one whose components can be realized via what we shall call a *standard Bell experiment*.³

Definition 6. A *standard Bell experiment* (in relation to $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$) on a source characterized by some quantum state ρ is one whereby the source distributes pairs of physical systems to Alice and Bob, and where each of them can perform (on each physical system that they receive), respectively, m_A and m_B alternative measurements that would each generate n_A and n_B distinct outcomes.

Here, we have implicitly assumed that at the receiving ends, the composite systems that Alice and Bob receive are still well characterized by the same physical state ρ and this is the assumption that we will make whenever we deal with a standard Bell experiment. With this assumption, then via a standard Bell experiment, the sets of local POVM elements $\{\{A_{s_a}^{o_a}\}_{o_a=1}^{n_A}\}_{s_a=1}^{m_A}$, $\{\{B_{s_b}^{o_b}\}_{o_b=1}^{n_B}\}_{s_b=1}^{m_B}$ and the (bipartite) quantum state ρ give rise to a quantum correlation $\mathbf{p}_{\text{QM}} \in \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ via Eq. (4.2). As such, we will also say that ρ , together with these POVM elements form a *quantum strategy* that realizes \mathbf{p}_{QM} .

Of course, at a more general level, one can also imagine a scenario where Alice and Bob choose to perform local measurements on $N > 1$ copies of the quantum systems at a time; this is the scenario of performing a standard Bell experiment on $\rho^{\otimes N}$. Alternatively, one could also imagine that while the source is well characterized by ρ , Alice and Bob may choose to perform some local preprocessing on ρ which effectively transforms it to some other state ρ' prior to a standard Bell experiment. Loosely, we shall say that these are nonstandard Bell experiments on ρ , since the source is still well characterized by ρ . However, these and other scenarios which do not fit within the framework of a standard Bell experiment on ρ will be the topics of future discussion in Chapter 6 and Chapter 7.

4.2.1 General Structure of the Set of Quantum Correlations

Now, let us take a closer look at the structure of $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, and in particular its relationship with $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$. To begin with, we note that for any two quantum probability vectors \mathbf{p}_{QM} , $\mathbf{p}'_{\text{QM}} \in \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, an arbitrary convex combination of them

$$\mathbf{p}'' = q \mathbf{p}_{\text{QM}} + (1 - q) \mathbf{p}'_{\text{QM}}, \quad (4.3)$$

where $0 \leq q \leq 1$ also gives rise to another quantum probability vector $\mathbf{p}'' \in \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$. To see this, let us denote by ρ and $\{\{A_{s_a}^{o_a}\}_{o_a=1}^{n_A}\}_{s_a=1}^{m_A}$, $\{\{B_{s_b}^{o_b}\}_{o_b=1}^{n_B}\}_{s_b=1}^{m_B}$, respectively, a quantum

³In the literature, the term *standard Bell experiment* has been used in various different contexts. In particular, it is commonly used to refer to a Bell experiment that involves measurements of two-dichotomic observables per site (i.e., $m_A = m_B = n_A = n_B = 2$). When there are only two parties involved in the experiment, this reduces to an experiment that tests against the Bell-CHSH/ Bell-CH inequality. Here, we are using this term in the same sense as that used in Ref. [147], which distinguishes it from nonstandard Bell experiment that typically involves (either active or passive) preprocessing prior to an actual Bell test.

state and some local POVM which together form a quantum strategy for \mathbf{p}_{QM} ; likewise, ρ' , $\{\{A'_{s_a}\}_{o_a=1}^{n_A}\}_{s_a=1}^{m_A}$ and $\{\{B'_{s_b}\}_{o_b=1}^{n_B}\}_{s_b=1}^{m_B}$ which together realize the quantum probability vector \mathbf{p}'_{QM} . Then it is easy to see that the quantum state

$$\rho'' \equiv q \rho \oplus (1 - q) \rho', \quad (4.4)$$

and the local POVM (elements) defined by

$$A''_{s_a} \equiv A_{s_a}^{o_a} \oplus A'_{s_a}{}^{o_a}, \quad B''_{s_b} \equiv B_{s_b}^{o_b} \oplus B'_{s_b}{}^{o_b}, \quad (4.5)$$

for all o_a, o_b, s_a and s_b do realize the probability vector \mathbf{p}'' in the sense of Eq. (4.2). Hence, as with the set of classical correlations, $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ is convex. However, in sharp contrast with $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, the set of quantum correlations is *not* a convex polytope [62]. Nonetheless, for the simplest scenario where $m_A = m_B = n_A = n_B = 2$, the boundary of the set of quantum correlations, or more precisely $\mathcal{Q}_{2;2}^{2;2}$ has already been characterized in Refs. [144, 145].

How is $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ related to $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$? Intuitively, one would expect $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ to be a subset of $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$. To see that this is indeed the case, it suffices to show that all extreme points of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ are contained in $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, that is, all extremal classical probability vectors can be realized by some quantum strategy. For definiteness, let us consider the extreme point $\mathbf{a}, \mathbf{b} \mathbf{B}_{AB}$ whose components are given by Eq. (3.10). A particular trivial way to realize this classical probability vector is to pick any (normalized) quantum state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ and the following local POVM elements

$$A_{s_a}^{o_a} = \delta_{o_a} \vartheta_{s_a}^{[1]} \mathbf{1}_{d_A}, \quad B_{s_b}^{o_b} = \delta_{o_b} \vartheta_{s_b}^{[2]} \mathbf{1}_{d_B}, \quad (4.6)$$

for all o_a, o_b, s_a and s_b . Then, from Eq. (4.2), it is straightforward to see that this quantum strategy does realize the classical probabilities given in Eq. (3.10). Therefore, $\text{vert}(\mathcal{P}_{m_A; m_B}^{n_A; n_B}) \subset \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ and by convexity of $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$ and $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, it follows that $\mathcal{P}_{m_A; m_B}^{n_A; n_B} \subseteq \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, i.e., the set of classical correlations is contained in the set of quantum correlations.

On the other hand, as we recall from Bell's theorem (Theorem 2), there are quantum correlations which violate a Bell inequality and hence fall outside the set of classical correlations (hereafter we will also refer to a probability vector \mathbf{p} which is in $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ but not in $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$ as a *nonclassical correlation*). Therefore, the set of quantum correlations $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ is a strict superset of the set of classical correlations $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, i.e., $\mathcal{P}_{m_A; m_B}^{n_A; n_B} \subset \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$, for at least some choices of m_A, m_B, n_A and n_B . Meanwhile, it has also been known for some time that the set of quantum correlations $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ is a strict subset of the set of correlations satisfying the no-signaling condition, i.e., $\mathcal{Q}_{m_A; m_B}^{n_A; n_B} \subset \mathcal{C}_{m_A; m_B}^{n_A; n_B}$ [103]. In fact, some quantitative understanding on the volume of these three sets, namely, $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$, $\mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ and $\mathcal{C}_{m_A; m_B}^{n_A; n_B}$ has recently been established for the simplest scenario of $m_A = m_B = n_A = n_B = 2$ [148].

4.2.2 Quantum Correlation and Bell Inequality Violation

Although all quantum states are capable of generating classical correlations, only some quantum states are capable of generating correlations outside the classical correlation polytope. Necessarily, in this case, the nonclassical correlation $\mathbf{p}_{\text{QM}} \in \mathcal{Q}_{m_A; m_B}^{n_A; n_B}$ must give rise to a violation of some Bell inequality. As we shall see in the later chapters, the kind of correlation

that a quantum state ρ can offer depends very much on whether it is a standard or a non-standard Bell experiment that is carried out on ρ . However, even if we restrict our attention to standard Bell experiments, c.f. Definition 6, whether a given quantum state can offer nonclassical correlation may still depend on the actual number of possible measurements — m_A and m_B — as well as the actual number of possible outcomes for each measurement — n_A and n_B (see Chapter 5 and 6 for examples).

In this regard, let us now introduce the following definition for a Bell inequality violation by a given state ρ with respect to some specific choice of the parameters $\mathbf{m} \equiv (m_A, m_B)$ and $\mathbf{n} \equiv (n_A, n_B)$.

Definition 7. A quantum state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is said to violate a Bell inequality $I_{\mathbf{m};\mathbf{n}}^{(k)}$, Eq. (3.12) – Eq. (3.14), via a standard Bell experiment if and only if \exists local measurement operators $\{\{A_{s_a}^{o_a}\}_{o_a=1}^{n_A}\}_{s_a=1}^{m_A} \subset \mathcal{B}(\mathcal{H}_A)$, $\{\{B_{s_b}^{o_b}\}_{o_b=1}^{n_B}\}_{s_b=1}^{m_B} \subset \mathcal{B}(\mathcal{H}_B)$ such that the resulting quantum probability vector \mathbf{p}_{QM} obtained via Eq. (4.2) violates $I_{\mathbf{m};\mathbf{n}}^{(k)}$.

Hereafter, unless otherwise stated, Bell inequality violation for a given state ρ will always be used in relation to a standard Bell experiment and with respect to some specific Bell inequality $I_{\mathbf{m};\mathbf{n}}^{(k)}$. As we shall see later in Sec. 4.3, it is possible that ρ *does not violate* any Bell inequalities or is *known to satisfy* a large class of Bell inequalities for all choices of local measurements. At this stage, it is worth noting that, as with the set of quantum states, the set of quantum states that do not violate a given Bell inequality is *convex* (see Appendix B.2.1 for a proof).

Nevertheless, as long as ρ *does violate* a Bell inequality with some choice of local measurements, we will say that ρ is Bell-inequality-violating:

Definition 8. A quantum state ρ is said to be Bell-inequality-violating (henceforth abbreviated as BIV) if and only if for some \mathbf{m} and \mathbf{n} , ρ violates a Bell inequality $I_{\mathbf{m};\mathbf{n}}^{(k)}$ for some k in the sense defined in Definition 7. Similarly, a quantum state ρ is said to be non-Bell-inequality-violating (henceforth abbreviated as NBIV) if and only if for all \mathbf{m} and \mathbf{n} , ρ does not violate any Bell inequality in the sense defined in Definition 7.⁴

Clearly, since the set of quantum states not violating a specific Bell inequality is convex, so is the set of quantum states that are NBIV. Let us denote this set by \mathcal{NV} . As far as a standard Bell experiment is concerned, the behavior of NBIV quantum states is entirely classical, since any experimental statistics generated from these states can be mimicked by some LHVM. In the next section, we will review some well-known examples of quantum states which are NBIV as well as quantum states which are known to satisfy a large class of Bell inequalities.

4.3 Locally Causal Quantum States

Historically, Bell inequality violation has served as one of the first means, both theoretical and experimental, to demonstrate stronger than classical correlations. Nevertheless, as is

⁴A quantum state that is BIV is commonly known in the literature as a *nonlocal* state; likewise, a quantum state that is NBIV is commonly known in the literature as a *local* state. This convention, however, is not unanimously accepted (see, for example, Ref. [149]).

now well-known, not all quantum states are capable of demonstrating such nonclassical correlations. In fact, some quantum states are *only* capable of generating classical correlations in *any* (standard) Bell experiments.

4.3.1 Separable States

An obvious example of a quantum state that is only capable of producing classical correlations is a *separable state* (aka a *classically correlated state* [29]). In its simplest form, an n -partite separable pure state $|\Psi_{\text{sep}}\rangle$ is just the tensor product of n pure states, i.e.,

$$|\Psi_{\text{sep}}\rangle = |\phi^{[1]}\rangle \otimes |\phi^{[2]}\rangle \otimes \cdots \otimes |\phi^{[n]}\rangle, \quad (4.7)$$

where $|\phi^{[i]}\rangle \in \mathcal{H}^{[i]}$. Due to its form, these separable states are also known as *product states*, or sometimes *uncorrelated states*. It is easy to see that measurement statistics on a single particle, say that described by $|\phi^{[i]}\rangle$, can be modeled in a purely classical manner. In particular, the state vector $|\phi^{[i]}\rangle$, or more generally $|\Psi_{\text{sep}}\rangle$, serves as a perfectly legitimate LHV that reproduces the quantum mechanical predictions. This can be seen, for example, by noting that the joint probability of observing some local measurement outcomes always factorizes into the n marginal probabilities and thus the correlation generated is always classical, c.f. Eq. (3.7).

Of course, separable states can also be correlated. The most general separable state involves one that can be decomposed as a convex combination of product states, i.e.,

$$\rho_{\text{sep}} = \sum_k p_k \rho_k^{[1]} \otimes \rho_k^{[2]} \otimes \cdots \otimes \rho_k^{[n]}, \quad (4.8a)$$

$$p_k \geq 0, \quad \sum_k p_k = 1, \quad (4.8b)$$

where $\rho_k^{[i]} \equiv |\phi_k^{[i]}\rangle\langle\phi_k^{[i]}|$. Operationally, these are states that can be prepared from classical correlations using only local quantum operations assisted by classical communication (henceforth abbreviated as LOCC⁵) [29]. Since each term in the sum, i.e., $\otimes_{i=1}^n \rho_k^{[i]}$ is NBIV, and hence can be modeled classically, so is their convex combination. An immediate consequence of this is thus the following Lemma [29].

Lemma 9. *A quantum state describing a composite system is BIV only if it is entangled, i.e., non-separable⁶ across its subsystems [29]. Hence, a BIV state cannot be written in the form of Eq. (4.8).*

4.3.2 Quantum States Admitting General LHV

Naively, it seems plausible that the converse of Lemma 9, i.e., “all entangled states are BIV”, is true. In other words, for any given entangled state, there exist appropriate measurements such that the corresponding correlation vector obtained from Eq. (4.2) lies outside

⁵This is also commonly known in the literature as LQCC (see, for example Ref. [150] and references therein).

⁶A non-separable state was originally called an EPR correlated state in Ref. [29].

the classical correlation polytope. However, it turns out that there are entangled quantum states whose measurement statistics in a standard Bell experiment can be reproduced entirely using some LHVM. In this section, we will look at some specific examples of quantum states whereby a general LHVM can be constructed to reproduce the quantum mechanical prediction for projective and/ or POVM measurements on the respective quantum states.

4.3.2.1 $U \otimes U$ Invariant States — Werner States

The first counterexample to the commonly held intuition that “entanglement \Rightarrow Bell inequality violation” was given by Werner [29] who considered bipartite quantum states $\rho_{\text{W}_d} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ that are invariant under $U \otimes U$, i.e.,

$$\rho_{\text{W}_d} = U \otimes U \rho_{\text{W}_d} U^\dagger \otimes U^\dagger, \quad (4.9)$$

where U is an arbitrary unitary operator acting on \mathbb{C}^d . It can be shown that ρ_{W_d} , now known as the Werner state, admits the following compact form⁷

$$\rho_{\text{W}_d}(q) = (1 - q) \frac{\Pi_+}{\text{tr}(\Pi_+)} + q \frac{\Pi_-}{\text{tr}(\Pi_-)}, \quad 0 \leq q \leq 1, \quad (4.11)$$

where Π_+ and Π_- are, respectively, the projector onto the symmetric and antisymmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$. Using the identities $\Pi_+ + \Pi_- = \mathbf{1}_d \otimes \mathbf{1}_d$, $\text{tr}(\Pi_\pm) = d(d \pm 1)/2$, Werner states can also be written as an affine combination of the antisymmetric projector and the $d \times d$ -dimensional maximally mixed state, i.e.,

$$\rho_{\text{W}_d}(p) = p \frac{2 \Pi_-}{d(d-1)} + (1 - p) \frac{\mathbf{1}_d \otimes \mathbf{1}_d}{d^2}, \quad 1 - \frac{2d}{d+1} \leq p \leq 1, \quad (4.12)$$

where $p = 1 - \frac{2d}{d+1}(1 - q)$.⁸ The separability of Werner states has been fully characterized in Ref. [29]: a Werner state is separable if and only if $p \leq p_{\text{s},\text{W}_d} \equiv 1/(d+1)$ or equivalently, $q \leq \frac{1}{2}$.

For any von Neumann (projective) measurement on $\rho_{\text{W}_d}(p)$ with

$$p = p_{\text{L},\text{W}_d}^\Pi \equiv 1 - \frac{1}{d}, \quad (4.13)$$

Werner [29] has constructed an LHVM that reproduces the corresponding quantum mechanical prediction.⁹ Recall that the set of quantum states not violating a given Bell inequality is convex (c.f. Appendix B.2.1), therefore for $d \geq 2$, $\rho_{\text{W}_d}(p)$ with $p_{\text{s},\text{W}_d} < p \leq p_{\text{L},\text{W}_d}^\Pi$ is entangled but does not violate any Bell inequality with projective measurements. Since $p_{\text{L},\text{W}_d}^\Pi$ is an

⁷Note that Werner has used, instead, the following parametrization in Ref. [29]:

$$\rho_{\text{W}_d}(\Phi) = \frac{1}{d^3 - d} [(d - \Phi) \mathbf{1}_{d^2} + (d\Phi - 1)V], \quad -1 \leq \Phi \leq 1, \quad (4.10)$$

where $V \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is the flip operator such that $V|\alpha\rangle_A |\beta\rangle_B = |\beta\rangle_A |\alpha\rangle_B$ and $\Phi = 1 - 2q$.

⁸Note that in this case, the weight p could also take on negative values.

⁹It is worth noting that alternative derivations of Werner’s LHVM for the $d = 2$ case could also be found in Refs. [32, 151].

increasing function of d , an interesting feature of Werner's model is that it covers a greater range of entangled Werner states as d increases (Figure 4.1).

Given that this is not the most general measurement that one can perform on $\rho_{w_d}(p)$, one may be tempted to conjecture that all entangled Werner states could produce nonclassical correlations with generalized measurements given by POVMs. In 2002, Barrett showed that this line of thought is untenable [31]. In particular, he constructed a LHV for Werner states with

$$p = p_{L, w_d}^{\text{POVM}} \equiv \frac{3d-1}{d^2-1} \left(1 - \frac{1}{d}\right)^d, \quad (4.14)$$

for any POVM measurement. It is not difficult to show that $p_{s, w_d} \leq p_{L, w_d}^{\text{POVM}}$ for any $d \geq 2$, thus from the convexity of NBIV states, it follows that any Werner state with $p_{s, w_d} \leq p \leq p_{L, w_d}^{\text{POVM}}$ is entangled but does not violate *any* Bell inequality. In contrast with Werner's model [29], p_{L, w_d}^{POVM} decreases as d increases, hence the applicability of Barrett's LHV shrinks as d increases (Figure 4.1). In fact, it can be easily checked that in the asymptotic limit of $d \rightarrow \infty$, Barrett's model is barely applicable to any entangled Werner states.

For the specific case of $d = 2$, since the antisymmetric projector Π_- is none other than the projector onto the Bell singlet state $|\Psi^-\rangle$, a two-qubit¹⁰ Werner state is essentially a noisy Bell singlet state. Building on earlier work by Tsirelson [145], Acín *et al.* [153] showed that in this case, i.e., $d = 2$, there exists an LHV for projective measurements on Werner states with $p \lesssim 0.65950$. This is done by first showing that the threshold p whereby $\rho_{w_2}(p)$ becomes NBIV with projective measurements, denoted by $p_{L, w_2}^{\Pi, c}$, is related to the Grothendieck's constant of order three, i.e., $K_G(3)$ by $p_{L, w_2}^{\Pi, c} = 1/K_G(3)$. Then, by using an upper bound¹¹ on $K_G(3)$ due to Krivine [154], the above lower bound on $p_{L, w_2}^{\Pi, c}$ follows immediately. Moreover, if any one of the observers has his/ her projective measurements restricted to a plane in the Bloch sphere, then there is a LHV for $\rho_{w_2}(p)$ if and only if $p \leq 1/\sqrt{2}$ [153, 154].

4.3.2.2 $U \otimes \bar{U}$ Invariant States — Isotropic States

Recently, a similar construction of an LHV was also obtained by Almeida *et al.* for another class of bipartite mixed states with a high degree of symmetry [33]. Isotropic states, as they are now known, were first introduced in Ref. [155] and have the nice property of being invariant under $U \otimes \bar{U}$, i.e.,

$$\rho_{I_d}(p) = U \otimes \bar{U} \rho_{I_d} U^\dagger \otimes \bar{U}^\dagger, \quad (4.15)$$

where \bar{U} denotes the complex conjugate of an arbitrary $d \times d$ unitary matrix U . As for Werner states, the isotropic states $\rho_{I_d} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ admit the explicit form [155]

$$\rho_{I_d}(p) = p |\Phi_d^+\rangle\langle\Phi_d^+| + (1-p) \frac{\mathbf{1}_d \otimes \mathbf{1}_d}{d^2}, \quad (4.16)$$

¹⁰A qubit is a two-level quantum system, which can be physically realized, for example, by the polarization of a photon, the spin of an electron etc. (see, for example, Ref. [152]). A two-qubit state, in this context, is the state of a two-party system, where each subsystem can be represented by a qubit.

¹¹The exact value of $K_G(3)$ is not known.

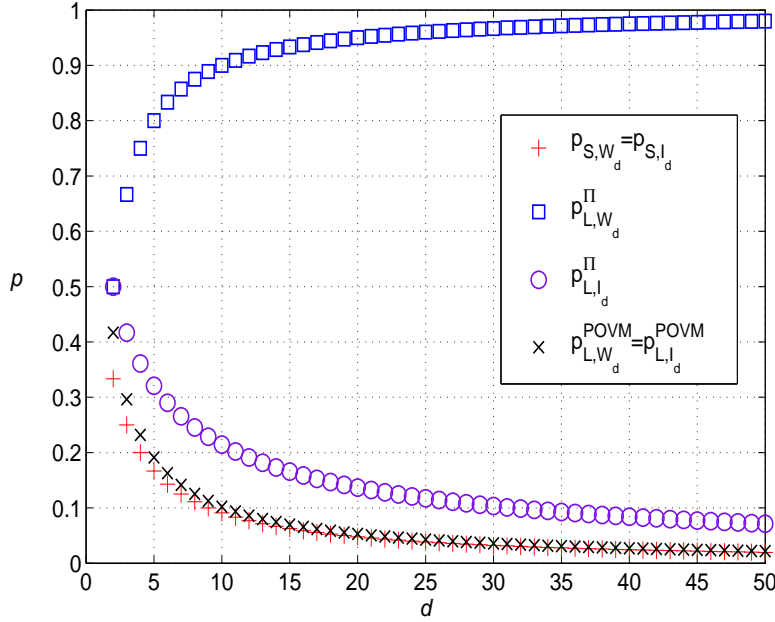


FIGURE 4.1: Plot of the various threshold weights p_{S,W_d} , p_{L,W_d}^{Π} , p_{L,W_d}^{POVM} for Werner states $\rho_{W_d}(p)$, and p_{S,I_d} , p_{L,I_d}^{Π} , p_{L,I_d}^{POVM} for isotropic states $\rho_{I_d}(p)$ as a function of d . Notice that for these two classes of states, the threshold weights for separability, i.e., p_{S,W_d} and p_{S,I_d} are identical; likewise for the threshold weights whereby a LHV for POVM measurements is known to exist, i.e., p_{L,W_d}^{POVM} and p_{L,I_d}^{POVM} . For each d , the vertical line joining these two threshold weights, which are, respectively, marked by a red + and a black ×, correspond to weights p of $\rho_{W_d}(p)$ and $\rho_{I_d}(p)$ whereby the states are entangled but do not violate any Bell inequalities. Similarly, for each d , the vertical line joining a blue □ and a red + corresponds to $\rho_{W_d}(p)$ which are entangled but are NBIV with projective measurements whereas the vertical line joining a purple circle and a red + corresponds to $\rho_{I_d}(p)$ that are entangled but are NBIV with projective measurements.

which, for $0 \leq p \leq 1$, can be interpreted as a convex mixture of the maximally entangled state $|\Phi_d^+\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$,

$$|\Phi_d^+\rangle \equiv \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle_A \otimes |i\rangle_B, \quad (4.17)$$

and the maximally mixed state, where $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ are, respectively, local orthonormal bases of $\mathcal{H}_A = \mathbb{C}^d$ and $\mathcal{H}_B = \mathbb{C}^d$.¹² If, and only if $p \leq p_{S,I_d} \equiv 1/(d+1)$, the mixture represents a separable state [155]. It is worth noting that in this case, partial transposition of $\rho_{I_d}(p)$ gives a legitimate and separable $\rho_{W_d}(p')$ with $p' = (1-d)p$ [15].

In the same spirit as Werner's and Barrett's construction, Almeida *et al.* constructed an

¹²Due to the explicit form given in Eq. (4.17), some authors also refer to the isotropic state $\rho_{I_d}(p)$ as the (generalized) Werner states (see, for example, Ref. [156]).

explicit LHVM for projective measurements as well as an LHVM for generalized measurements on $\rho_{I_d}(p)$. Specifically, their models work for mixtures with weight p given by

$$p_{L,I_d}^{\Pi} \equiv \frac{1}{d-1} \sum_{k=2}^d \frac{1}{k} \quad \text{and} \quad p_{L,I_d}^{\text{POVM}} \equiv \frac{3d-1}{d^2-1} \left(1 - \frac{1}{d}\right)^d = p_{L,W_d}^{\text{POVM}} \quad (4.18)$$

respectively.¹³ Just like p_{L,W_d}^{POVM} , p_{L,I_d}^{Π} is monotonically decreasing with d . Nevertheless, for any $d \geq 1$, it can again be shown that the latter is always greater than or equal to p_{S,I_d} (Figure 4.1). Hence, experimental statistics obtained from projective measurements on isotropic states with $p_{S,I_d} < p \leq p_{L,I_d}^{\Pi}$ cannot violate any Bell inequalities. More generally, isotropic states with $p_{S,I_d} < p \leq p_{L,I_d}^{\text{POVM}}$ are entangled but are NBIV.

Notice that when $d = 2$, the isotropic state is local unitarily equivalent to a Werner state. Therefore, all the bounds obtained by Acín *et al.* [153] for 2-dimensional Werner state are also applicable to the 2-dimensional isotropic state. In particular, this means that $p_{L,W_d}^{\Pi,c} = p_{L,I_d}^{\Pi,c}$. On the other hand, if only traceless observables are measured on the isotropic states, it was also shown in Ref. [153] that a LHVM for the experimental statistics exists for $p \leq 1/K_G(d^2 - 1)$, where $K_G(n)$ is the Grothendieck constant of order n .

4.3.2.3 $U \otimes U \otimes U$ Invariant States

In the multipartite scenario, Werner's LHVM has also been extended to cover some tripartite states with $U \otimes U \otimes U$ symmetry [158]. In Ref. [32], Tóth and Acín gave an alternative derivation of Werner's LHVM, which allows them to generalize straightforwardly to tripartite states of the form

$$\rho_{\text{TA}}(p) = \frac{1}{8} \mathbb{1}_2 \otimes \mathbb{1}_2 \otimes \mathbb{1}_2 + \frac{1}{24} \sum_{k=x,y,z} \mathbb{1}_2 \otimes \sigma_k \otimes \sigma_k - \frac{p}{16} (\sigma_k \otimes \mathbb{1}_2 \otimes \sigma_k + \sigma_k \otimes \sigma_k \otimes \mathbb{1}_2), \quad (4.19)$$

where $\{\sigma_i\}_{i=1}^3$ are Pauli matrices introduced in Eq. (2.6). Their model works for projective measurements and $p \leq 1$ whereas $\rho_{\text{TA}}(p)$ with $p > p_{S,\text{TA}} \equiv \frac{1}{3}(\sqrt{13} - 1)$ are states that cannot be written either in the form of Eq. (4.8) or the form

$$\begin{aligned} & \sum_k p_k^{[AB]} \rho_k^{[AB]} \otimes \rho_k^{[C]} + p_k^{[AC]} \rho_k^{[AC]} \otimes \rho_k^{[B]} + p_k^{[BC]} \rho_k^{[BC]} \otimes \rho_k^{[A]}, \\ & p_k^{[ij]} \geq 0 \quad \forall i, j \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}, \quad \sum_k p_k^{[AB]} + p_k^{[AC]} + p_k^{[BC]} = 1, \end{aligned} \quad (4.20)$$

where $\rho_k^{[i]} \in \mathcal{B}(\mathcal{H}_i)$, and $\rho_k^{[ij]} \in \mathcal{B}(\mathcal{H}_i \otimes \mathcal{H}_j)$ for all $i, j \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$. Tripartite states that can be written in the form of Eq. (4.20) are biseparable [159] and can be prepared by mixing pure states on one side and an entangled two-party state at the remaining sites. Hence, $\rho_{\text{TA}}(p)$ with $p_{S,\text{TA}} < p \leq 1$ contains genuine tripartite entanglement but admits LHVM for projective measurements.

¹³It is interesting to note that, from here, Almeida *et al.* [33] have also found, using existing results from Ref. [157], a lower bound on p whereby an arbitrary convex mixture of a bipartite pure state and the maximally mixed state would admit a LHVM for both projective and the generalized measurements.

4.3.3 Quantum States Satisfying Some Bell Inequalities

The LHVMs that have been constructed for $\rho_{w_d}(p)$, $\rho_{l_d}(p)$ and $\rho_{TA}(p)$ are very general in that they can reproduce exactly the quantum mechanical prediction for projective and/or POVM measurements on the respective states. As a result, these states do not violate *any* Bell inequalities via measurements where the models are applicable. However, construction of these general LHVMs are by no means trivial, and could only be done, so far, for states with a high-degree of symmetry. In this section, we will look at some examples of entangled states that are known to satisfy a large class of, rather than *all* Bell inequalities. For the examples presented in Sec. 4.3.3.1, no explicit LHVM is constructed, but the states are known to satisfy a large class of Bell inequalities whereas for the examples presented in Sec. 4.3.3.2, an LHVM is constructed for Bell inequalities with a specific number of measurement settings per site.

4.3.3.1 PPT Entangled States

Historically, positive-partial-transposed (henceforth abbreviated as PPT) entangled states referred to bipartite entangled states that remain positive semidefinite (henceforth abbreviated as PSD) after partial transposition with respect to one of its subsystems [160]. By virtue of this property, entanglement of PPT states cannot be decided using the Peres-Horodecki criterion (*aka* PPT criterion) for separability [161, 162]. The very first example of a PPT entangled state in the literature is the following 1-parameter family of two-qutrit¹⁴ mixed states [160]:

$$\rho_H(p) = \frac{8p}{8p+1}\rho_{\text{Ent}} + \frac{1}{8p+1}|\Psi_p\rangle\langle\Psi_p|, \quad 0 < p < 1, \quad (4.21a)$$

where

$$\rho_{\text{Ent}} = \frac{1}{8} \sum_{i,j=0,i \neq j}^2 |i\rangle\langle i| \otimes |j\rangle\langle j| - \frac{1}{8}|2\rangle\langle 2| \otimes |0\rangle\langle 0| + \frac{3}{8}|\Phi_3^+\rangle\langle\Phi_3^+|, \quad (4.21b)$$

$$|\Psi_p\rangle = |2\rangle \otimes \left(\sqrt{\frac{1+p}{2}}|0\rangle + \sqrt{\frac{1-p}{2}}|2\rangle \right), \quad (4.21c)$$

and $|\Phi_3^+\rangle$ is the maximally entangled state for $d = 3$, c.f. Eq. (4.17).

As was first demonstrated by Horodecki *et al.* [163], a bipartite PPT state cannot be distilled [164] to a Bell singlet state $|\Psi^-\rangle$ using LOCC. Hence, PPT entangled states are also known as *bound entangled states*. The entanglement contained in a bound entangled state is rather weak and often has to be used in conjunction with other entangled states to demonstrate its nonclassical features. In fact, it was even conjectured by Peres [8] that no PPT entangled states violate any Bell inequalities. The first result that was in favor of this conjecture was given by Werner and Wolf [165] where they showed, using the variance inequality,¹⁵ that an n -partite (entangled) state that is PPT with respect to all combinations of its subsystems cannot violate any of the n -partite Mermin inequalities, Eq. (3.44) –

¹⁴A qutrit is a three-level quantum system.

¹⁵That is, the variance of a random variable is non-negative.

Eq. (3.48). Since the Mermin inequality reduces to the Bell-CHSH inequality when $n = 2$, an immediate corollary of Werner and Wolf's result is that no bipartite PPT entangled states can violate the Bell-CHSH inequality.

It is still possible, however, to see a Bell inequality violation coming from an n -partite entangled state $\rho \in \mathcal{B}(\mathcal{H}^{[1]} \otimes \mathcal{H}^{[2]} \otimes \dots \mathcal{H}^{[n]})$ that is PSD with respect to transposition of each individual subsystem, i.e.,

$$\rho^{T_k} \geq 0 \quad \forall \quad k \in \{1, 2, \dots, n\}, \quad (4.22)$$

where $(\cdot)^{T_k}$ denotes the partial transposition with respect to subsystem k . In particular, Dür [166] showed that the n -partite Mermin inequality (with $n \geq 8$) is violated by an n -partite entangled state that is of this sort. Specifically, the multipartite mixed entangled state that Dür considered reads:¹⁶

$$\rho_D = \frac{1}{n+1} \left(|\Psi_{\text{GHZ}}\rangle\langle\Psi_{\text{GHZ}}| + \frac{1}{2} \sum_{k=1}^n (|\Phi_{k,0}\rangle\langle\Phi_{k,0}| + |\Phi_{k,1}\rangle\langle\Phi_{k,1}|) \right), \quad (4.23)$$

where $|\Psi_{\text{GHZ}}\rangle \in \mathcal{H}^{[1]} \otimes \mathcal{H}^{[2]} \otimes \dots \otimes \mathcal{H}^{[n]} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2$ is the n -partite generalized GHZ state [167],

$$|\Psi_{\text{GHZ}}\rangle \equiv \frac{1}{\sqrt{2}} (|0\rangle^{\otimes n} + e^{i\alpha_n} |1\rangle^{\otimes n}), \quad (4.24)$$

α_n is an arbitrary phase factor, $\{|\Phi_{k,0}\rangle\langle\Phi_{k,0}|, |\Phi_{k,1}\rangle\langle\Phi_{k,1}|\}$ are product states defined by

$$\begin{aligned} |\Phi_{k,0}\rangle &\equiv |0^{[1]}\rangle \otimes |0^{[2]}\rangle \otimes \dots \otimes |0^{[k-1]}\rangle \otimes |1^{[k]}\rangle \otimes |0^{[k+1]}\rangle \otimes \dots \otimes |0^{[n]}\rangle, \\ |\Phi_{k,1}\rangle &\equiv |1^{[1]}\rangle \otimes |1^{[2]}\rangle \otimes \dots \otimes |1^{[k-1]}\rangle \otimes |0^{[k]}\rangle \otimes |1^{[k+1]}\rangle \otimes \dots \otimes |1^{[n]}\rangle, \end{aligned}$$

and $\{|0^{[j]}\rangle, |1^{[j]}\rangle\}$ are local orthonormal basis vectors for $\mathcal{H}^{[j]}$. This is, nevertheless, not in contradiction with the result given by Werner and Wolf [165]. In fact, follow up work by Acín [168] showed that for all these states violating the Mermin inequality, there is at least one bipartite splitting of the system such that the state becomes distillable.¹⁷

Of course, as reviewed earlier in Sec. 3.3.3, Mermin inequality is not the only class of tight Bell correlation inequalities for the n -partite correlation polytope ${}^s\mathcal{P}_{2;2;\dots;2}^{2;2;\dots;2}$, thus a natural question that follows is whether Werner and Wolf's result [165] generalizes to all the 2^{2^n} tight Bell correlation inequalities with two dichotomic observables per site, c.f. Eq. (3.48) and Eq. (3.49). In 2001, Werner and Wolf [115] provided a positive answer to this question — n -partite states that are PPT with respect to all combinations of its subsystems do not violate any of the n -partite Bell correlation inequalities with two dichotomic observables per site. By far, this is the strongest result in support of Peres' conjecture. Although a counterexample to this conjecture is not known in the literature, the same goes for a proof, despite the wide range of supporting evidence. In what follows, we will review some other examples of bipartite PPT entangled states which are known to satisfy a large class of Bell inequalities.

¹⁶For $n \geq 4$, ρ_D has positive partial transposition with respect to each of the subsystem k , where $k \in \{1, 2, \dots, n\}$ but the state is not PSD if a partial transposition is carried out with respect to $\mathcal{H}^{[i]} \otimes \mathcal{H}^{[j]}$ for any $i \neq j$. Hence, by the PPT criterion for separability [161, 162], ρ_D for $n \geq 4$ cannot be fully separable, i.e., cannot be written in the form of Eq. (4.8).

¹⁷See also Ref. [169] for a more thorough discussion between distillability and violation of n -partite Bell inequality.

4.3.3.2 Entangled States with Symmetric Quasiextension

Given that it is nontrivial to come up with a general LHV, a natural question that follows is whether there is any systematic way to generate, perhaps not as general, LHV for arbitrary quantum states. In 2003, an important breakthrough along this line came about following Terhal *et al.*'s consideration of symmetric quasiextension for multipartite quantum states [108]. To appreciate that, let us recall the following definition from Ref. [108]:

Definition 10. Let $\pi : \mathcal{H}^{\otimes s} \rightarrow \mathcal{H}^{\otimes s}$ be a permutation of Hilbert spaces \mathcal{H} in $\mathcal{H}^{\otimes s}$ and let

$$\text{Sym}_{\mathcal{H}^{\otimes s}}(\rho) \equiv \frac{1}{s!} \sum_{\pi} \pi \rho \pi^\dagger, \quad (4.25)$$

then ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ has a (s_a, s_b) -symmetric quasiextension when there exists a multipartite entanglement witness¹⁸ $W_\rho \in \mathcal{B}(\mathcal{H}_A^{\otimes s_a} \otimes \mathcal{H}_B^{\otimes s_b})$ such that $\text{tr}_{\mathcal{H}_A^{\otimes(s_a-1)} \otimes \mathcal{H}_B^{\otimes(s_b-1)}} W_\rho = \rho$ and $W_\rho = \text{Sym}_{\mathcal{H}_A^{\otimes s_a}} \otimes \text{Sym}_{\mathcal{H}_B^{\otimes s_b}}(W_\rho)$.

With this definition of symmetric quasiextension, Terhal *et al.* [108] then went on to show that if ρ has an (s_a, s_b) -symmetric quasiextension, then an LHV can be constructed for ρ for all Bell experiments with $\mathbf{m} = (s_a, s_b)$; hence, ρ does not violate any Bell inequality with $\mathbf{m} = (s_a, s_b)$ settings. In fact, the following strengthened version of the theorem was also proven in the same paper [108].

Theorem 11. If ρ has a $(1, s_b)$ -symmetric quasiextension, then ρ does not violate a Bell inequality with s_b settings for Bob and any number of settings for Alice. Similarly, if ρ has a $(s_a, 1)$ -symmetric quasiextension, then ρ does not violate a Bell inequality with s_a settings for Alice and any number of settings for Bob.

From Definition 10, it follows that if a given state has an (s_a, s_b) -symmetric quasiextension, it must necessarily have a $(1, s_b)$ -symmetric quasiextension and an $(s_a, 1)$ -symmetric quasiextension. This, together with Theorem 11, implies that if ρ has an (s_a, s_b) -symmetric quasiextension, it cannot violate any Bell inequalities with $\mathbf{m} = (s_{a'}, s_{b'})$ settings where $\min\{s_{a'}, s_{b'}\} \leq \max\{s_a, s_b\}$. As a first application of their technique, Terhal *et al.* constructed a $(2, 2)$ -symmetric extension for any bipartite bound entangled state based on a real *unextendible product basis* [170]. Therefore, if any of such states is to violate a Bell inequality, it must involve more than 2 measurement settings on at least one of the sites.

The construction of a symmetric (quasi)extension of a given quantum state ρ , if it exists, can be done, to some extent, numerically. In particular, the search for an (s_a, s_b) -symmetric quasiextension with non-negative¹⁹ or decomposable W_ρ is a semidefinite programming feasibility problem (Appendix C.3.3), which can be efficiently solved on a computer. In some

¹⁸An n -partite entanglement witness W is a Hermitian matrix that satisfies

$$\text{tr}(W \rho_{\text{Sep}}) \geq 0$$

for all n -partite separable states ρ_{Sep} , c.f. Eq. (4.8).

¹⁹In this case, the corresponding entanglement witness W_ρ is a trivial one and it actually corresponds to what is called a symmetric extension of ρ [171, 172].

cases, these semidefinite programs (henceforth abbreviated as SDP) can even be solved analytically. For example, in the case of Werner states, c.f. Eq. (4.12), it was established in Ref. [108, 173] that all $\rho_{W_d}(p)$ have symmetric extensions as long as

$$s_a + s_b \leq d. \quad (4.26)$$

In the case of $d = 2$, a bound better than Eq. (4.26) was also derived in Ref. [173], namely, a $(2, 2)$ -symmetric extension and hence an LHV with 2 settings or less can be constructed for $\rho_{W_2}(p)$ with $-\frac{1}{3} \leq p \leq \frac{2}{3}$.²⁰ Note that, in comparison with the work presented by Acín *et al.* [153], the LHV derived in this manner is actually applicable to more entangled $\rho_{W_2}(p)$ even for POVM measurements. The tradeoff, however, is that it is only applicable to scenarios where $\min\{s_a, s_b\} \leq 2$.

Numerically, Terhal *et al.*'s construction has also been applied to the following one-parameter family of two-qutrit mixed state [174]:

$$\rho_{\text{CH}}(\alpha) = \frac{2}{7}|\Phi_3^+\rangle\langle\Phi_3^+| + \frac{\alpha}{7}\sigma_+ + \frac{5-\alpha}{7}\sigma_-, \quad 2 < \alpha < 5, \quad (4.27a)$$

where

$$\sigma_{\pm} = \frac{1}{3} \sum_{j=0}^2 |j\rangle\langle j| \otimes |j \pm 1 \bmod 3\rangle\langle j \pm 1 \bmod 3|, \quad (4.27b)$$

which is known to be separable for $2 \leq \alpha \leq 3$, bound entangled for $3 < \alpha \leq 4$ and having negative partial transposition for $4 < \alpha \leq 5$. In particular, entangled $\rho_{\text{CH}}(\alpha)$ was found to possess a $(2, 2)$ -symmetric quasiextension and a $(3, 3)$ -symmetric quasiextension derived from a decomposable entanglement witness for $\alpha \in [3, 4.84]$ and $\alpha \in [3, 4]$ respectively. Therefore any potential Bell inequality violation of the bound entangled $\rho_{\text{CH}}(\alpha)$ must involve at least four alternative measurements on one of the sites.

4.4 Conclusion

In this chapter, we have formally defined what we mean by quantum correlations, and the closely related concept of a standard Bell experiment. We have also looked at some of the basic structure of the set of quantum correlations and its relationship with the classical correlation polytope. In addition, we have also reviewed some well-known examples of quantum states admitting either a partial, or a full LHV for projective/ POVM measurements. The stage is finally set for us to look into genuine quantum correlations which cannot be accounted for by any LHV.

²⁰It should be emphasized that these bounds were obtained by considering a symmetric quasiextension derived from either a non-negative entanglement witness or a decomposable entanglement witness. It could very well be that the state of interest has a symmetric quasiextension that is not of either of these two forms.

5

Bounds on Quantum Correlations in Standard Bell Experiments

As we have seen in the previous chapter, correlations generated by quantum systems can sometimes be described in a purely classical manner via a local hidden variable model. By Bell's theorem, of course, we know that some entangled quantum states can also offer correlations that are not describable within the classical framework. In this and the next chapter, we will look at such nonclassical behavior displayed by entangled quantum systems in standard Bell experiments.

5.1 Introduction

Before pursuing any in-depth study on the nonclassical correlations offered by quantum systems, it seems natural to first determine if a given entangled state is Bell-inequality-violating (BIV) and hence capable of demonstrating nonclassical correlations in a standard Bell experiment. In the terminologies that we have introduced earlier in Sec. 3.2.2, this amounts to determining if a given quantum state, with a judicious choice of local measurements, can give rise to correlations that lie outside the classical correlation polytope. Typically, this is done by varying over the local measurements that each observer may perform and checking if the resulting statistics can violate any Bell inequalities.

Surprisingly, relatively little is known in terms of which quantum states are BIV. For bipartite quantum systems, the strongest results that we know in this regard are due to Gisin and Peres [36], who showed that all bipartite pure entangled states violate the Bell-CHSH inequality (a weaker version of Gisin and Peres's result was first presented by Capasso *et al.* [34] and later rediscovered by Gisin [35]). In other words, a bipartite pure quantum system is capable of demonstrating nonclassical correlations *if and only if* it is entangled.

The corresponding situation for multipartite quantum systems is a lot more complicated

and it is still not known if all multipartite pure entangled states are BIV. To begin with, Scarani and Gisin [175] noticed that some generalized GHZ states, despite being entangled, do not violate any of the Mermin inequalities, Eq. (3.44) – Eq. (3.48). Although some of these states were later found to violate some, among the complete set of 2^{2^n} n -partite correlation inequalities, Eq. (3.48) – Eq. (3.49), the rest were proved to satisfy this set of inequalities also [176]. A twist came about when Chen *et al.* [37] constructed a tripartite Bell inequality for probabilities and proved that all the above-mentioned generalized GHZ states, as well as any 2-entangled pure tripartite states¹ violate the constructed Bell inequality. In addition, they have presented some numerical evidence that this inequality is also violated by other kinds of tripartite entangled pure states. For $N > 3$ parties, some further investigations exist (see Refs. [139, 141, 177] and references therein) but nothing as strong as the results presented by Gisin and Peres in Ref. [36] is known yet.

As for mixed quantum states, Horodecki *et al.* have also provided an analytic criterion [30] to determine if a two-qubit state violates the Bell-CHSH inequality. This criterion is, unfortunately, also the only analytic criterion that we have in determining if a broad class of quantum states, namely two-qubit states, can be simulated by some LHV in a standard Bell experiment. Nonetheless, for specific quantum states, such as those that we have looked at in Chapter 4, the existence of LHVMs for these states will exclude the possibility of them violating a Bell inequality (via measurements where the models are applicable).

In general, to determine if a quantum state violates a Bell inequality is a high-dimensional variational problem, which requires a nontrivial optimization of a Hermitian operator \mathcal{B} (now known as the *Bell* operator [178]) over the various possible measurement settings that each observer may perform. This optimization does not appear to be convex and is possibly NP-hard [179]. In fact, a closely related problem, namely to determine if a given probability vector is a member of the set of classical correlations is known to be NP-complete [62].

Except for the simplest scenario where one deals with the Bell-CHSH inequality, in conjunction with a two-qubit state [30], or a (bipartite) maximally entangled pure state [36, 180], and its mixture with the maximally mixed state [14], very few analytic results for the optimal measurements are known. As such, for the purpose of characterizing quantum states that are incompatible with locally causal description, efficient algorithms to perform this state-dependent optimization are very desirable.

On the other hand, state-independent bounds of quantum correlations have also been investigated since the early 1980s. In particular, Tsirelson [143] has demonstrated, using what is now known as Tsirelson’s vector construction, that in a Bell-CHSH setup, bipartite quantum systems of arbitrary dimensions cannot exhibit correlations stronger than $2\sqrt{2}$ - a value now known as Tsirelson’s bound. Recently, analogous bounds for more complicated Bell inequalities have also been investigated by Filipp and Svozil [54], Buhrman and Massar [55], Wehner [56], Toner [57], Avis *et al.* [58] and Navascués *et al.* [17]. On a related note, bounds on quantum correlations for given local measurements, rather than given quantum state, have also been investigated by Cabello [181] and Bovino *et al.* [182].

¹These are tripartite states of the form

$$|\Psi_{ABC}\rangle = |\Psi_{AB}\rangle \otimes |\Psi_C\rangle, \quad (5.1)$$

where $|\Psi_{AB}\rangle$ is a bipartite pure entangled state.

The main purpose of this chapter is to look into the algorithmic aspect of determining if a quantum state can violate a given Bell inequality. In particular, we will present, respectively, in Sec. 5.2.2 and Sec. 5.2.3, two algorithms that were developed to provide a lower bound and an upper bound on the maximal expectation value of a Bell operator for a given quantum state. The second algorithm is another instance where a nonlinear optimization problem is approximated by a hierarchy of semidefinite programs, each giving a better bound of the original optimization problem [17, 172, 183, 184, 185]. In its simplest form, it provides a bound that is apparently state-independent.

In Sec. 5.3.1, we will derive, based on the second algorithm, a necessary condition for a class of two-qudit states² to violate the Bell-CHSH inequality. Next, in Sec. 5.3.2, we will illustrate how the lower bound algorithm can be used to derive the Horodecki criterion [30] for two-qubit states. After that, we will demonstrate how the two algorithms can be used in tandem to determine if some quantum states violate a given Bell inequality. Some limitations of these algorithms will then be discussed. We will conclude with a summary of results and some possibilities for future research.

5.2 Bounds on Quantum Correlations

5.2.1 Preliminaries

In the earlier chapter, we have learned that a particular Bell inequality deals with a specific experimental setup, say involving two experimenters Alice and Bob,³ where each of them can perform, respectively, m_A and m_B alternative measurements that would each generate n_A and n_B distinct outcomes. For each of these setups, a Bell inequality places a bound on the experimental statistics obtained from the corresponding Bell experiments. In particular, we recall from Sec. 3.2.3 that a (linear) Bell inequality takes the form :

$$\mathcal{S}_{\text{LHV}} \leq \beta_{\text{LHV}}, \quad (5.2)$$

where β_{LHV} is a real number and \mathcal{S}_{LHV} involves a specific *linear* combination of correlation functions or joint and marginal probabilities of experimental outcomes.

To determine if a quantum state violates a given Bell inequality with some choice of measurements, we need to evaluate these correlation functions, or probabilities according to the quantum mechanical rules [see Eq. (2.3) for an example]. The bounds on \mathcal{S}_{LHV} then translate into corresponding bounds β_{LHV} on the expectation value of some Hermitian observable that describes the (standard) Bell inequality experiment, this observable is known as the *Bell operator* \mathcal{B} [178]. The restriction that the given Bell inequality is satisfied in the experiment is then

$$\mathcal{S}_{\text{QM}}(\rho, \mathcal{B}) = \text{tr}(\rho \mathcal{B}) \leq \beta_{\text{LHV}}. \quad (5.3)$$

The Bell operator depends on the choice of measurements at each of the sites (polarizer angles for example). These measurements will be described by a set of Hermitian operators

²A two-qudit state is a bipartite quantum state describing two d -level quantum systems. Some authors refer to them, instead, as a two-qunit state for two n -level quantum systems.

³For definiteness, we will restrict our attention to bipartite setups and point out, when relevant, how the arguments can be extended to the multipartite scenario.

$\{O_m\}$. For correlation inequalities these are simply the measured observables at each stage of the Bell measurement, while for general probability inequalities the O_m are POVM elements that describe the measurements at each site. We will denote this expectation value by $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ when we want to emphasize its dependence on the choice of local Hermitian observables O_m . Ideally the choice of measurement should give the maximal expectation value of the Bell operator, for which we will give the notation

$$\mathcal{S}_{\text{QM}}(\rho) \equiv \max_{\{O_m\}} \mathcal{S}_{\text{QM}}(\rho, \{O_m\}). \quad (5.4)$$

Moreover, we will explicitly include a superscript in $\mathcal{S}_{\text{QM}}(\rho)$, e.g. $\mathcal{S}_{\text{QM}}^{(k)}(\rho)$, when we want to make reference to a specific Bell inequality labeled by “ k ”. It is this implicitly-defined function that will give us information about which states violate a given Bell inequality.

As an example, let us recall the Bell-CHSH inequality, Eq. (2.26), which is reproduced here for ease of reference

$$\mathcal{S}_{\text{LHV}}^{(\text{CHSH})} = E(A_1, B_1) + E(A_1, B_2) + E(A_2, B_1) - E(A_2, B_2) \leq 2. \quad (5.5)$$

In quantum mechanics, each of these correlation functions $E(A_{s_a}, B_{s_b})$ is computed using

$$E_{\text{QM}}(A_{s_a}, B_{s_b}) = \text{tr}(\rho A_{s_a} \otimes B_{s_b}). \quad (5.6)$$

Substituting this into Eq. (5.5) and comparing with Eq. (5.3), one finds that the corresponding Bell operator reads

$$\mathcal{B}_{\text{CHSH}} = A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2). \quad (5.7)$$

To determine the maximal Bell-inequality violation for a given ρ , $\mathcal{S}_{\text{QM}}(\rho)$, requires a maximization by varying over all possible choices of $\{O_m\}$, i.e., A_{s_a} and B_{s_b} in the case of Eq. (5.7). Whether we are interested in correlation inequalities or in Bell inequalities for probabilities the (bipartite) Bell operator has the general structure

$$\mathcal{B} = \sum_{K,L} b_{KL} A_K \otimes B_L, \quad (5.8)$$

which essentially follows from the linearity of Bell inequality as well as the linearity of expectation values in quantum mechanics. In the case of a Bell inequality for probabilities the indices K, L are collective indices, c.f. Eq. (3.16), describing both a particular measurement setting and a particular outcome for each observer; the A_K and B_L are then POVM elements corresponding to specific outcomes in the Bell experiment. For correlation inequalities, the indices K, L refer simply to the measurement settings as in the Bell-CHSH case described in detail above.

In what follows, we will present two algorithms which we have developed specifically to perform the maximization over the choice of measurements. The first, which we will abbreviate as LB, provides a *lower bound* on the maximal expectation value and can be implemented for any Bell inequality. This bound makes use of the fact that the objective function $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ is bilinear in the observables O_m , that is it is linear in the A_K for fixed

B_L and likewise linear in the B_L for fixed A_K . The second bound, which we will abbreviate as UB, provides an *upper bound* on $\mathcal{S}_{\text{QM}}(\rho)$ by regarding $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ as a polynomial function of the variables that define the various O_m and applying general techniques for finding such bounds on polynomials [183, 184]. Both of these make use of convex optimization techniques in the form of a semidefinite program (SDP). An SDP is a linear optimization over positive semidefinite (PSD) matrices which are subjected to affine constraints. Readers who are unfamiliar with semidefinite programming are referred to Appendix C.1.

5.2.2 Algorithm to Determine a Lower Bound on $\mathcal{S}_{\text{QM}}(\rho)$

The key idea behind the LB algorithm is to realize that when measurements for all but one party are *fixed*, the optimal measurements for the remaining party can be obtained efficiently using convex optimization techniques, in particular an SDP. Thus we can fix Bob's measurements and find Alice's optimal choice, at least numerically; with these optimized measurements for Alice, we can further find the optimal measurements for Bob (for this choice of Alice's settings), and then Alice again and so on and so forth until $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ converges within the desired numerical precision.⁴

Back in 2001, Werner and Wolf [104] presented a similar iterative algorithm, by the name of *See-Saw iteration*, to maximize the expectation value of the Bell operator for a Bell correlation inequality involving only dichotomic observables.⁵ As a result we will focus here on the (straightforward) generalization to the widest possible class of Bell inequalities. In the work of Werner and Wolf [104] it turned out that once the dichotomic observables for one party are fixed, optimization of the other party's observables can be carried out explicitly. This turns out to be true for any dichotomic Bell inequality and we will return to this question in Sec. 5.2.2.3.

5.2.2.1 General Settings

Let us now consider a Bell inequality for probabilities for $\mathcal{P}_{m_A; m_B}^{n_A; n_B}$.⁶ We will denote the POVM element associated with the o_a^{th} outcome of Alice's s_a^{th} measurement by $A_{s_a}^{o_a}$ while $B_{s_b}^{o_b}$ is the POVM element associated with the o_b^{th} outcome of Bob's s_b^{th} measurement. Moreover, let d_A and d_B , respectively, be the dimension of the state space that each of the $A_{s_a}^{o_a}$ and $B_{s_b}^{o_b}$ acts on. Then it follows from Born's rule that

$$p_{AB}^{o_a o_b}(s_a, s_b) = \text{tr}(\rho A_{s_a}^{o_a} \otimes B_{s_b}^{o_b}) \quad (4.2a)$$

$$p_A^{o_a}(s_a) = \text{tr}(\rho A_{s_a}^{o_a} \otimes \mathbf{1}_{d_B}), \quad p_B^{o_b}(s_b) = \text{tr}(\rho \mathbf{1}_{d_A} \otimes B_{s_b}^{o_b}), \quad (4.2b)$$

where, as defined in Sec. 2.2.2, $p_{AB}^{o_a o_b}(s_a, s_b)$ refers to the joint probability that the o_a^{th} experimental outcome is observed at Alice's site and the o_b^{th} outcome at Bob's, given that Alice performs the s_a^{th} measurement and Bob performs the s_b^{th} measurement; likewise for the

⁴That such an iterative algorithm using SDP can lead to a local maximum of $\mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}, B_{s_b}^{o_b})$ was also discovered independently by Ito *et al.* [14].

⁵A dichotomic observable is a Hermitian observable with only two distinct eigenvalues.

⁶For a Bell correlation inequality, we can apply LB by first rewriting the corresponding Bell operator, Eq. (5.16), in terms of POVM elements that form the measurement outcomes of each O_m .

marginal probabilities $p_A^{o_a}(s_a)$ and $p_B^{o_b}(s_b)$. A general Bell operator for probabilities can then be expressed as

$$\mathcal{B} = \sum_{s_a=1}^{m_A} \sum_{o_a=1}^{n_A} \sum_{s_b=1}^{m_B} \sum_{o_b=1}^{n_B} b_{s_a s_b}^{o_a o_b} A_{s_a}^{o_a} \otimes B_{s_b}^{o_b}, \quad (5.10)$$

where $b_{s_a s_b}^{o_a o_b}$ are determined from the given Bell inequality, c.f. Eq. (3.16). Again, one is reminded that the sets of POVM elements $\{A_{s_a}^{o_a}\}_{o_a=1}^{n_A}$ and $\{B_{s_b}^{o_b}\}_{o_b=1}^{n_B}$ satisfy

$$\sum_{o_a=1}^{n_A} A_{s_a}^{o_a} = \mathbb{1}_{d_A} \quad \text{and} \quad \sum_{o_b=1}^{n_B} B_{s_b}^{o_b} = \mathbb{1}_{d_B} \quad \forall \quad s_a, s_b, \quad (4.1a)$$

$$A_{s_a}^{o_a} \geq 0, \quad B_{s_b}^{o_b} \geq 0 \quad \forall \quad s_a, s_b, o_a, o_b. \quad (4.1b)$$

5.2.2.2 Iterative Semidefinite Programming Algorithm

To see how to develop a lower bound on $\mathcal{S}_{\text{QM}}(\rho)$ by fixing the observables at one site and optimizing the other, we observe that upon substituting Eq. (5.10) into Eq. (5.3), the *lhs* of the inequality can be rewritten as

$$\mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}, B_{s_b}^{o_b}) = \sum_{s_b, o_b} \text{tr} \left(\rho_{B_{s_b}^{o_b}} B_{s_b}^{o_b} \right), \quad (5.12)$$

where

$$\rho_{B_{s_b}^{o_b}} \equiv \sum_{s_a, o_a} b_{s_a s_b}^{o_a o_b} \text{tr}_A \left[\rho \left(A_{s_a}^{o_a} \otimes \mathbb{1}_{d_B} \right) \right], \quad (5.13)$$

and $\text{tr}_A \cdot$ is the partial trace over subsystem \mathcal{A} .

Notice that if all $\rho_{B_{s_b}^{o_b}}$ are held constant by fixing all of Alice's measurement settings (given by the set of $A_{s_a}^{o_a}$) then $\rho_{B_{s_b}^{o_b}}$ is a constant matrix independent of the $B_{s_b}^{o_b}$. Thus the objective function is linear in these variables. The constraints that $\{B_{s_b}^{o_b}\}_{o_b=1}^{n_B}$ form a POVM for each value of s_b is a combination of affine and matrix nonnegativity constraints. As a result it is fairly clear that the following problem is an SDP in *standard form*, Eq. (C.1),

$$\text{maximize}_{\{B_{s_b}^{o_b}\}} \quad \mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}, B_{s_b}^{o_b}) \quad (5.14a)$$

$$\text{subject to} \quad \sum_{o_b=1}^{n_B} B_{s_b}^{o_b} = \mathbb{1}_{d_B} \quad \forall \quad s_b, \quad (5.14b)$$

$$B_{s_b}^{o_b} \geq 0 \quad \forall \quad s_b, o_b. \quad (5.14c)$$

The detailed formulation of this optimization problem in terms of an SDP in standard form can be found in Appendix C.3.1.

Exactly the same analysis follows if we fix Bob's measurement settings, and optimize over Alice's POVM elements instead. To arrive at a local maximum of $\mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}, B_{s_b}^{o_b})$, it therefore suffices to start with some random measurement settings for Alice (or Bob), and optimize over the two parties' settings iteratively. A (nontrivial) lower bound on $\mathcal{S}_{\text{QM}}(\rho)$ can then be obtained by optimizing the measurement settings starting from a set of randomly generated initial guesses.

It is worth noting that in any implementation of this algorithm, physical observables $\{A_{s_a}^{o_a}, B_{s_b}^{o_b}\}$ achieving the lower bound are constructed when the corresponding SDP is solved. In the event that the lower bound is greater than the classical threshold β_{LHV} , then these observables can, in principle, be measured in the laboratory to demonstrate a Bell-inequality violation of the given quantum state.

We have implemented this algorithm in MATLAB⁷ to search for a lower bound on $\mathcal{S}_{\text{QM}}(\rho)$ in the case of Bell-CH, I_{3322} , I_{4422} , I_{2233} and I_{2244} inequalities (Sec. 3.3.1), and with the local dimension $d = d_{\mathcal{A}} = d_{\mathcal{B}}$ up to 32. Typically, with no more than 50 iterations, the algorithm already converges to a point that is different from a local maximum by no more than 10^{-9} . To test the effectiveness of finding $\mathcal{S}_{\text{QM}}(\rho)$ using LB, we have randomly generated 200 Bell-CH violating two-qubit states and found that on average, it takes about 6 *random initial guesses* before the algorithm gives $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ that is close to the actual maximum, computed using Horodecki's criterion [30], to within 10^{-5} . Specific examples of the implementation of this algorithm will be discussed in Sec. 5.3 and Chapter 6.

Two other remarks concerning this algorithm should now be made. Firstly, the algorithm is readily generalized to multipartite Bell inequalities for probabilities: one again starts with some random measurement settings for all but one party, and optimizes over each party iteratively. Also, it is worth noting that this algorithm is not only useful as a numerical tool, but for specific cases, it can also provide a useful analytic criterion. In particular, when applied to the Bell-CH inequality for two-qubit states, the LB algorithm may lead us to the Horodecki criterion [30], i.e., the necessary and sufficient condition for two-qubit states to violate the Bell-CH/ Bell-CHSH inequality (see Sec. 5.3.2 and Appendix B.3.2 for more details).

5.2.2.3 Two-outcome Bell Experiment

We will show that, just as in the case of Bell correlation inequalities [104], the local optimization can be solved analytically for two-outcome measurements. If we denote by “ \pm ” the two outcomes of the experiments, it follows from Eq. (4.1) that the POVM element $B_{s_b}^-$ can be expressed as a function of the complementary POVM element $B_{s_b}^+$, i.e., $B_{s_b}^- = \mathbf{1}_{d_{\mathcal{B}}} - B_{s_b}^+$, subjected to $0 \leq B_{s_b}^+ \leq \mathbf{1}_{d_{\mathcal{B}}}$. We then have

$$\sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} B_{s_b}^{o_b} \right) = \text{tr} \left[\left(\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-} \right) B_{s_b}^+ \right] + \text{tr} \left(\rho_{B_{s_b}^-} \right).$$

The above expression can be maximized by setting the PSD operator $B_{s_b}^+$ to be the projector onto the positive eigenspace of $\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-}$. In a similar manner, we can also write

$$\sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} B_{s_b}^{o_b} \right) = \text{tr} \left[\left(\rho_{B_{s_b}^-} - \rho_{B_{s_b}^+} \right) B_{s_b}^- \right] + \text{tr} \left(\rho_{B_{s_b}^+} \right),$$

which can be maximized by setting $B_{s_b}^-$ to be the projector onto the non-positive eigenspace of $\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-}$. Notice that this choice is consistent with our earlier choice of $B_{s_b}^+$ for the “+”

⁷MATLAB is a trademark of The Math Works, Inc., Natick, MA.

outcome POVM element in that they form a valid POVM. Since there can be no difference in these maxima, we may write the maximum as their average, i.e.,

$$\sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} B_{s_b}^{o_b} \right) = \frac{1}{2} \left\| \rho_{B_{s_b}^+} - \rho_{B_{s_b}^-} \right\| + \frac{1}{2} \sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} \right),$$

where $\|O\|$ is the trace norm of the Hermitian operator O [186, 187]. Carrying out the optimization for each of the m_B settings, the optimized $\mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}, B_{s_b}^{o_b})$, as an implicit function of Alice's POVM $\{A_{s_a}^{o_a}\}$, is given by

$$\mathcal{S}_{\text{QM}}(\rho, A_{s_a}^{o_a}) = \frac{1}{2} \sum_{s_b} \left\| \rho_{B_{s_b}^+} - \rho_{B_{s_b}^-} \right\| + \frac{1}{2} \sum_{s_b} \sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} \right). \quad (5.15)$$

Notice that this calculation is essentially the same as that which shows that the Helstrom measurement [188] is optimal for distinguishing two quantum states.

An immediate corollary of the above result is that for the optimization of a two-outcome Bell operator for probabilities, it is unnecessary for any of the two observers to perform generalized measurements described by a POVM; von Neumann projective measurements are sufficient.⁸ In practice, this simplifies any analytic treatment of the optimization problem as a generic parametrization of a POVM is a lot more difficult to deal with, thereby supporting the simplification adopted in Ref. [54].

Nevertheless, it may still be advantageous to consider generic POVMs as our initial measurement settings when implementing the algorithm numerically. This is because the local maximum of $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ obtained using the iterative procedure is a function of the initial guess. In particular, it was found that the set of local maxima attainable could change significantly if the ranks of the initial measurement projectors are altered. As such, it seems necessary to step through various ranks of the starting projectors to obtain a good lower bound on $\mathcal{S}_{\text{QM}}(\rho)$. Even then, we have also found examples where this does not give a lower bound on $\mathcal{S}_{\text{QM}}(\rho)$ that is as good as when generic POVMs are used as the initial measurement operators.

5.2.3 Algorithm to Determine an Upper Bound on $\mathcal{S}_{\text{QM}}(\rho)$

A major drawback of the above algorithm, or the analogous algorithm developed by Werner and Wolf [104] for Bell correlation inequalities is that, except in some special cases, it is generally impossible to tell if the maximal $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ obtained through this optimization procedure corresponds to the global maximum $\mathcal{S}_{\text{QM}}(\rho)$.

Nontrivial upper bounds on $\mathcal{S}_{\text{QM}}(\rho)$, nevertheless, can be obtained by considering *relaxations* of the global optimization problem given by Eq. (5.4). In a relaxation, a (possibly nonconvex) maximization problem is modified in some way so as to yield a more tractable optimization that bounds the optimization of interest. One example of a variational upper bound that exists for any optimization problem is the *Lagrange dual* optimization that arises in the method of Lagrange multipliers [63].

⁸As was pointed out in Ref. [14], this sufficiency also follows from Theorem 5.4 in Ref. [189].

To see how to apply existing studies in the optimization literature to find upper bounds on $\mathcal{S}_{\text{QM}}(\rho)$, let us first remark that the global objective function $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$ can be mapped to a polynomial function in real variables, for instance, by expanding all the local observables $\{O_m\}$ and the density matrix ρ in terms of Hermitian basis operators. In the same manner, matrix equality constraints, such as that given in Eq. (4.1a) can also be mapped to a set of polynomial equalities by requiring that the matrix equality holds component wise. Now, it is known from the work of Lasserre [184] and Parrilo [183] that a hierarchy of global bounds of a polynomial function, subjected to polynomial equalities and inequalities, can be achieved by solving suitable SDPs. Essentially, this is achieved by approximating the original nonconvex optimization problem by a series of convex ones in the form of a SDP, each giving a better bound of the original polynomial objective function.

At the bottom of this hierarchy is the lowest order relaxation provided by the *Lagrange dual* of the original nonconvex problem. By considering *Lagrange multipliers* that depend on the original optimization variables, higher order relaxations to the original problem can be constructed to give tighter upper bounds on $\mathcal{S}_{\text{QM}}(\rho)$ (see Appendix C.2 for more details).

In the following, we will focus our discussion on a general two-outcome Bell (correlation) inequality, where the observables $\{O_m\}$ are only subjected to matrix equalities. In particular, we will show that the global optimization problem for these Bell inequalities is a *quadratically-constrained quadratic-program* (QCQP), i.e., one whereby the objective function and the constraints are both *quadratic* in the optimization variables. Then, we will demonstrate explicitly how the Lagrange dual of this QCQP, which is known to be an SDP, can be constructed. The analogous analytic treatment is apparently formidable for higher order relaxations. Nonetheless, there exists third-party MATLAB toolbox known as the SOSTOOLS which is tailored specifically for this kind of optimization problem [190, 191].

Numerically, we have implemented the algorithm for several two-outcome correlation inequalities and will discuss the results in greater detail in Sec. 5.3. For a general Bell inequality where each O_m is also subjected to a linear matrix inequality (henceforth abbreviated as LMI) like Eq. (4.1b), the algorithm can still be implemented, for instance, by requiring that all the principle submatrices of O_m have non-negative determinants [186, 187]. This then translates into a set of polynomial inequalities which fit into the framework of a general polynomial optimization problem (see Appendix C.2). However a more effective approach would retain the structure of linear matrix inequalities constraining a polynomial optimization problem; we leave the investigation of these bounds to further work.

5.2.3.1 Global Optimization Problem

Now, let us consider a dichotomic Bell correlation inequality where Alice and Bob can respectively perform m_A and m_B alternative measurements. A general Bell correlation operator

for such an experimental setup can be written as⁹

$$\mathcal{B} = \sum_{s_a=1}^{m_A} \sum_{s_b=1}^{m_B} b_{s_a s_b} O_{s_a} \otimes O_{s_b+m_A}, \quad (5.16)$$

where $b_{s_a s_b}$ are determined from the given Bell correlation inequality, O_{s_a} for $s_a = 1, \dots, m_A$ refers to the s_a^{th} Hermitian observable measured by Alice, and $O_{s_b+m_A}$ for $s_b = 1, \dots, m_B$ refers to the s_b^{th} Hermitian observable measured by Bob. Furthermore, these dichotomic observables are usually chosen to have eigenvalues ± 1 and thus

$$O_m^\dagger O_m = (O_m)^2 = \mathbf{1}_d \quad (5.17)$$

for all $m = 1, 2, \dots, m_A + m_B$, where we have assumed for simplicity that all the local observables O_m act on a state space of dimension d .¹⁰

The global optimization problem derived from a dichotomic Bell correlation inequality thus takes the form of

$$\text{maximize} \quad \text{tr}(\rho \mathcal{B}), \quad (5.18a)$$

$$\text{subject to} \quad O_m^2 = \mathbf{1}_d \quad (5.18b)$$

for all $m = 1, 2, \dots, m_A + m_B$. For any $m \times n$ complex matrices, we will now define $\text{vec}(A)$ to be the $m \cdot n$ dimensional vector obtained by stacking all columns of A on top of one another. By collecting all the vectorized observables together

$$\mathbf{w}^\dagger \equiv [\text{vec}(O_1)^\dagger \text{vec}(O_2)^\dagger \dots \text{vec}(O_{m_A+m_B})^\dagger],$$

and using the identity

$$\text{tr}(\rho O_{s_a} \otimes O_{s_b+m_A}) = \text{vec}(O_{s_a})^\dagger (V\rho)^{T_A} \text{vec}(O_{s_b+m_A}), \quad (5.19)$$

with V being the flip operator introduced in Eq. (4.10) and $(\cdot)^{T_A}$ being the partial transposition with respect to subsystem \mathcal{A} , we can write the objective function more explicitly as

$$\mathcal{S}_{\text{QM}}(\rho, \{O_m\}) = \text{tr}(\rho \mathcal{B}) = -\mathbf{w}^\dagger \Omega_0 \mathbf{w} \quad (5.20)$$

where

$$\Omega_0 \equiv \frac{1}{2} \begin{pmatrix} \mathbf{0} & -b \otimes R \\ -b^T \otimes R^\dagger & \mathbf{0} \end{pmatrix}, \quad (5.21)$$

b is a $m_A \times m_B$ matrix with $[b]_{s_a, s_b} = b_{s_a s_b}$, c.f. Eq. (5.16), and $R \equiv (V\rho)^{T_A}$. In this form, it is explicit that the objective function is *quadratic* in $\text{vec}(O_m)$. Similarly, by requiring that the

⁹Strictly, a general Bell correlation operator may also contain marginal terms like $O_m \otimes \mathbf{1}_{d_B}$ and $\mathbf{1}_{d_A} \otimes O_m$, that result from *restricted correlation function*, Eq. (3.38), defining the correlation inequality. For brevity, we will not consider such Bell operators in the following discussion. It should, nevertheless, be clear to the readers that the following arguments are readily generalized to include Bell operators of this more general kind.

¹⁰In general, the composite system may consist of subsystems of different dimensions. All the following arguments can be readily generalized to this more general scenario.

matrix equality, Eq. (5.17), holds component-wise, we can get a set of equality constraints, which are each *quadratic* in $\text{vec}(O_m)$. The global optimization problem given by Eq. (5.18) is thus an instance of a QCQP.

On a related note, for any Bell inequality experiments where measurements are restricted to the projective type, the global optimization problem is also a QCQP. To see this, we first note that the global objective function for the general case, as follows from Eq. (5.3) and Eq. (5.8), is always quadratic in the local Hermitian observables $\{A_K, B_L\}$. The requirement that these measurement operators are projectors amounts to requiring

$$A_K^2 = A_K, \quad B_L^2 = B_L, \quad \forall K, L, \quad (5.22)$$

which are quadratic constraints on the local Hermitian observables. Since we have shown in Sec. 5.2.2.3 that for any two-outcome Bell inequality for probabilities, it suffices to consider projective measurements in optimizing $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$, it follows that the global optimization problem for these Bell inequalities is always a QCQP.

5.2.3.2 State-independent Bound

As mentioned above, the lowest order relaxation to the global optimization problem — given by Eq. (5.18) — is simply the Lagrange dual of the original QCQP. This can be obtained via the *Lagrangian* [63] of the global optimization problem, i.e.,

$$\mathcal{L}(\{O_m\}, \Lambda_m) = \mathcal{S}_{\text{QM}}(\rho, \{O_m\}) - \sum_{m=1}^{m_A+m_B} \text{tr} [\Lambda_m (O_m^2 - \mathbf{1}_d)], \quad (5.23)$$

where Λ_m is a matrix of Lagrange multipliers associated with the m^{th} matrix equality constraint. With no loss of generality, we can assume that the Λ_m 's are Hermitian.

Notice that for all values of $\{O_m\}$ that satisfy the constraints, the Lagrangian

$$\mathcal{L}(\rho, \{O_m\}, \Lambda_m) = \mathcal{S}_{\text{QM}}(\rho, \{O_m\}).$$

As a result, if we maximize the Lagrangian without regard to the constraints we obtain an upper bound on the maximal expectation value of the Bell operator

$$\max_{\{O_m\}} \mathcal{L}(\rho, \{O_m\}, \Lambda_m) \geq \mathcal{S}_{\text{QM}}(\rho). \quad (5.24)$$

The Lagrange dual optimization simply looks for the best such upper bound.

In order to maximize the Lagrangian we rewrite the Lagrangian with Eq. (5.20) and the identity

$$\text{tr} (\Lambda_m O_m O_m^\dagger) = \text{vec}(O_m)^\dagger (\mathbf{1}_d \otimes \Lambda_m) \text{vec}(O_m), \quad (5.25)$$

to obtain

$$\mathcal{L}(\mathbf{w}, \Lambda_m) = -\mathbf{w}^\dagger \Omega \mathbf{w} + \sum_{m=1}^{m_A+m_B} \text{tr} \Lambda_m, \quad (5.26)$$

where

$$\Omega \equiv \Omega_0 + \bigoplus_{m=1}^{m_A+m_B} \mathbf{1}_d \otimes \Lambda_m. \quad (5.27)$$

Note that each of the diagonal blocks $\mathbf{1}_d \otimes \Lambda_m$ is of the same size as the matrix R .

To obtain the dual optimization problem, we maximize the Lagrangian over \mathbf{w} to obtain the *Lagrange dual function*

$$g(\Lambda_m) \equiv \sup_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \Lambda_m). \quad (5.28)$$

As noted above $g(\Lambda_m) \geq \mathcal{S}_{\text{QM}}(\rho)$ for all choices of Λ_m . Moreover, this supremum over \mathbf{w} is unbounded above unless $\Omega \geq 0$, in which case the supremum is attained by setting $\mathbf{w} = 0$ in Eq. (5.26). Hence, the *Lagrange dual* optimization, which seeks for the best upper bound of Eq. (5.18) by minimizing Eq (5.28) over the Lagrange multipliers, reads

$$\begin{aligned} & \text{minimize} && \sum_{m=1}^{m_A+m_B} \text{tr } \Lambda_m, \\ & \text{subject to} && \Omega \geq 0. \end{aligned} \quad (5.29)$$

By expanding Λ_m in terms of Hermitian basis operators satisfying Eq. (C.15),

$$\Lambda_m = \sum_{n=0}^{d^2-1} \lambda_{mn} \sigma_n, \quad (5.30)$$

the optimization problem given by Eq. (5.29) is readily seen to be an SDP in the *inequality form*, Eq. (C.2).

For Bell-CHSH inequality and the correlation equivalent of I_{3322} inequality given by Eq. (3.40), it was observed numerically that the upper bound obtained via the SDP (5.29) is always *state-independent*. For 1000 randomly generated two-qubit states, and 1000 randomly generated two-qutrit states, the Bell-CHSH upper bound of $\mathcal{S}_{\text{QM}}(\rho)$ obtained through (5.29) was never found to differ from the Tsirelson bound [143] by more than 10^{-7} . In fact by finding an explicit feasible solution to the optimization problem dual to Eq. (5.29), Wehner has shown that the upper bound obtained here can be no better than that obtained by Tsirelson's vector construction for correlation inequalities¹¹.

In a similar manner, we have also investigated the upper bound of $\mathcal{S}_{\text{QM}}(\rho)$ for some dichotomic Bell probability inequalities using the lowest order relaxation to the corresponding global optimization problem. Interestingly, the numerical upper bounds obtained from the analog of Eq. (5.29) for these inequalities – namely the Bell-CH inequality, Eq. (3.23), the I_{3322} inequality, Eq. (3.27), and the I_{4422} inequality, Eq. (3.28) — are also found to be state-independent and are given by 0.207 106 7, 0.375 and 0.669 346 1 respectively.

5.2.3.3 State-dependent Bound

Although the state-independent upper bounds obtained above are interesting in their own right, our main interest here is to find an upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ that does depend on the given quantum state ρ . This can be obtained, with not much extra cost, from the Lagrange dual to a more-refined version of the original optimization problem.

¹¹S. Wehner (private communication). See also Ref. [56].

To appreciate that, let us first recall that each dichotomic Hermitian observable O_m can only have eigenvalues ± 1 . It follows that their trace

$$z_m \equiv \text{tr}(O_m), \quad (5.31a)$$

can only take on the following values

$$z_m = -d, -d + 2, \dots, d - 2, d. \quad (5.31b)$$

In particular, if $z_m = \pm d$ for any m , then $O_m = \pm \mathbb{1}_d$ and it is known that the Bell-CHSH inequality cannot be violated for this choice of observable [178] (see also Appendix B.3.1).

Better Lagrange dual bounds arise from taking these additional constraints (5.31) explicitly into account. For that matter, we found it most convenient to express the original optimization problem in terms of real variables given by the expansion coefficients of O_m in terms of a basis for Hermitian matrices that includes the (traceless) Gell-Mann matrices and the identity matrix, c.f. Eq. (C.15). The resulting calculation is very similar to what we have done in the previous section (for details see Appendix C.2.1). Here, we will just note that the result is a set of SDPs, one for each of the various choices of z_m . The lowest order upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ can then be obtained by stepping through the various choices of z_m given in Eq. (5.31b), solving each of the corresponding SDPs, and taking their maximum. The results of this approach will be discussed later, for now it suffices to note that tighter bounds can be obtained that are explicitly state dependent.

5.2.3.4 Higher Order Relaxations

The higher order relaxations simply arise from allowing the Lagrange multipliers λ to be polynomial functions of $\{O_m\}$ rather than constants. In this case, it is no longer possible to optimize over the primal variables in the Lagrangian analytically but let us consider the following optimization

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && \gamma - \mathcal{S}_{\text{QM}}(\rho, \mathbf{x}) = \mu(\mathbf{x}) + \sum_i \lambda_i(\mathbf{x}) f_{\text{eq},i}(\mathbf{x}), \end{aligned} \quad (5.32)$$

where each $\lambda_i(\mathbf{x})$ is a polynomial function of \mathbf{x} and $\mu(\mathbf{x})$ is a sum of squares (SOS) polynomial and therefore non-negative. That is $\mu(\mathbf{x}) = \sum_j [h_j(\mathbf{x})]^2 \geq 0$ for some set of real polynomials $h_j(\mathbf{x})$. The variables of the optimizations are γ and the coefficients that define the polynomials $\mu(\mathbf{x})$ and $\lambda_i(\mathbf{x})$. Notice that we have $\gamma \geq \mathcal{S}_{\text{QM}}(\rho, \mathbf{x})$ whenever the constraints are satisfied so that once again we have a global upper bound on $\mathcal{S}_{\text{QM}}(\rho, \mathbf{x})$. This optimization can be implemented numerically by restricting $\mu(\mathbf{x})$ and $\lambda_i(\mathbf{x})$ to be of some fixed degree. The Lagrange dual optimization (5.29) arises from choosing the degree of $\lambda_i(\mathbf{x})$ to be zero. It is known that for any fixed degree this optimization is an SDP [183, 184] and we have implemented up to degree four using SOSTOOLS [190, 191]. Schmüdgen's theorem [192] guarantees that by increasing the degree of the polynomials in the relaxation we obtain bounds approaching the true maximum $\mathcal{S}_{\text{QM}}(\rho)$. This is a special case of the general procedure described in [183, 191, 192] which is also able to handle inequality constraints. For more details see Appendix C.2.

5.3 Applications & Limitations of the Two Algorithms

In this section, we will look at some concrete examples of how the two algorithms can be used to determine if some quantum states violate a Bell inequality. Specifically, we begin by looking at how the second algorithm can be used to determine, both numerically and analytically, if some bipartite qudit state violates the Bell-CHSH inequality. Then in Sec. 5.3.2, we will illustrate how LB can be used to recover the Horodecki criterion. After that, in Sec. 5.3.3, we demonstrate how the two algorithms can be used in tandem to determine if a class of two-qubit states violate the I_{3322} inequality, Eq. (3.27). We will conclude this section by pointing out some limitations of the UB algorithm that we have observed.

5.3.1 Bell-CHSH violation for Two-Qudit States

The Bell-CHSH inequality, as given by Eq. (5.5), is one that amounts to choosing [c.f. Eq. (5.16)]

$$b = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (5.33)$$

For low-dimensional quantum systems, an upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$ can be efficiently computed in MATLAB following the procedures described in Sec. 5.2.3.3. However, for high-dimensional quantum systems, intensive computational resources are required to compute this upper bound, which may render the computation infeasible in practice. In this regard, it is worth noting that for each choice of the Lagrange multipliers in the Lagrange dual function (5.28), there is a corresponding upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$. In fact, for a specific class of two-qudit states, namely those whose coherence vectors¹² vanish, and using some choice of the Lagrange multipliers, it can be shown (Appendix C.2.2) that $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$ cannot exceed

$$\max_{z_1, z_2, z_3, z_4} 2\sqrt{2}s_1 d \sqrt{\prod_{i=1}^2 \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{s_a, s_b=1}^2 b_{s_a s_b} \frac{z_{s_a} z_{s_b+2}}{d^2}, \quad (5.34)$$

where s_1 is the largest singular value of the matrix R' defined in Eq. (C.6b), and z_m is the trace of the dichotomic observable O_m given in Eq. (5.31). Since this bound is derived by considering a specific choice of the Lagrange multipliers, it is generally not as tight as the upper bound obtained numerically using the procedures described in Sec. 5.2.3.3.

To violate the Bell-CHSH inequality, we must have $\mathcal{S}_{\text{QM}}(\rho) > 2$, hence for this class of quantum states, the Bell-CHSH inequality cannot be violated if

$$\max_{z_1, z_2, z_3, z_4} \sqrt{2}s_1 d \sqrt{\prod_{i=1}^2 \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{s_a, s_b=1}^2 b_{s_a s_b} \frac{z_{s_a} z_{s_b+2}}{2d^2} \leq 1. \quad (5.35)$$

In addition, since the Bell-CHSH inequalities are the only class of nontrivial facet-inducing inequalities for ${}^c\mathcal{P}_{2;2}^{2;2}$, Eq. (5.35) guarantees the existence of an LHM for the experimental setup defined by $m_A = m_B = n_A = n_B = 2$ [99].

¹²These are generalization of the Bloch vectors representation for higher dimensional quantum systems. See also Refs. [193, 194].

Table 5.1: The various threshold values for isotropic states $\rho_{I_d}(p)$. The first column of the table is the dimension of the local subsystem d . From the second column to the seventh column, we have, respectively, the value of p below which the state is separable p_{S,I_d} ; the value of p below which Eq. (5.35) is satisfied $p_{\text{UB-semianalytic}}$, and hence the state does not violate the Bell-CHSH inequality; the value of p below which the upper bound obtained from lowest order relaxation is compatible with Bell-CHSH inequality; the value of p below which the state cannot violate any Bell inequality via projective measurements; the value of p below which the state cannot violate any Bell inequality (Sec. 4.3.2.2); and the value of p above which a Bell-CHSH violation has been observed using the LB algorithm.

d	p_{S,I_d}	$p_{\text{UB-semianalytic}}$	$p_{\text{UB-numerical}}$	p_{L,I_d}^{Π}	p_{L,I_d}^{POVM}	p_{LB}
2	0.33333	0.70711	0.70711	0.50000	0.41667	0.70711
3	0.25000	0.70711	0.70711	0.41667	0.29630	0.76297
4	0.20000	0.65465	0.65465	0.36111	0.23203	0.70711
5	0.16667	0.63246	0.63246	0.32083	0.19115	0.74340
10	0.09091	0.51450	-	0.21433	0.10214	0.70711
25	0.03846	0.36490	-	0.11733	0.04274	0.71516
50	0.01961	0.26963	-	0.07141	0.02171	0.70711

As an example, consider the d -dimensional isotropic state $\rho_{I_d}(p)$ introduced in Eq. (4.16). Recall from Sec. 4.3.2.2 that this class of states is entangled if and only if $p > p_{S,I_d} = 1/(d+1)$. Using the procedures outlined in Sec. 5.2.3.3, we can numerically compute, up till $d = 5$, the threshold value of p below which there can be no violation of the Bell-CHSH inequality; these critical values, denoted by $p_{\text{UB-numerical}}$ can be found in column 4 of Table 5.1. Similarly, we can numerically compute the corresponding threshold values given by Eq. (5.35), denoted by $p_{\text{UB-semianalytic}}$. It is worth noting that these threshold values, as can be seen from column 3 and 4 of Table 5.1, agree exceptionally well, thereby suggesting that the computationally feasible criterion given by Eq. (5.35) may be exact for the isotropic states.

5.3.2 Bell-CH violation for Two-Qubit States

The semianalytic criterion presented in Eq. (5.35) is general enough that it can be applied to any two-qudit states with vanishing coherence vectors. The price of such generality, however, is that the bound is often not tight. In particular, for $d = 2$, the exact value of $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$ for any two-qubit state ρ is known [30] and is often below the upper bound given by Eq. (5.34), i.e., $4\sqrt{2} s_1$.

Nevertheless, in this case, it turns out that we can use LB to obtain analytically $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$ and hence $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$ via

$$\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho) = 4 \left(\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) + \frac{1}{2} \right). \quad (5.36)$$

The Horodecki criterion [30] can then be recovered by imposing the condition $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) > 0$, or equivalently $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho) > 2$. To see this, let us first note that we can write a general

two-qubit state ρ in the so-called Hilbert-Schmidt form [30]¹³

$$\rho = \frac{1}{4} \left(\mathbf{1}_2 \otimes \mathbf{1}_2 + \mathbf{r}_A \cdot \vec{\sigma} \otimes \mathbf{1}_2 + \mathbf{1}_2 \otimes \mathbf{r}_B \cdot \vec{\sigma} + \sum_{i,j=x,y,z} [T]_{ij} \sigma_i \otimes \sigma_j \right), \quad (5.37)$$

where $\vec{\sigma}$ is defined in Eq. (2.5), $\{\sigma_i\}_{i=x,y,z}$ are the Pauli matrices introduced in Eq. (2.6) and

$$[T]_{ij} \equiv \text{tr}(\rho \sigma_i \otimes \sigma_j). \quad (5.38)$$

For ease of reference, we will now reproduce the functional form of the Bell-CH inequality as follows:

$$\mathcal{S}_{\text{LHV}}^{(\text{CH})} = p_{AB}^{o_a o_b}(1, 1) + p_{AB}^{o_a o_b}(1, 2) + p_{AB}^{o_a o_b}(2, 1) - p_{AB}^{o_a o_b}(2, 2) - p_A^{o_a}(1) - p_B^{o_b}(1) \leq 0. \quad (3.23)$$

where $o_a, o_b = \pm$ are the two possible local measurement outcomes in a two-outcome Bell-CH experiment. Substituting Eq. (4.2) into Eq. (3.23) and comparing the resulting expression with Eq. (5.3), one finds that the Bell operator for this Bell inequality with $o_a = o_b = +$ can be written as

$$\begin{aligned} \mathcal{B}_{\text{CH}} &= A_1^+ \otimes (B_1^+ + B_2^+) + A_2^+ \otimes (B_1^+ - B_2^+) - A_1^+ \otimes \mathbf{1}_{d_B} - \mathbf{1}_{d_A} \otimes B_1^+, \\ &= A_1^+ \otimes (B_2^+ - B_1^-) - A_2^- \otimes B_1^+ - A_2^+ \otimes B_2^+ \end{aligned} \quad (5.39)$$

where we have also made use of Eq. (4.1a) to arrive at the final form.¹⁴

Now, recall that it suffices to consider projective measurements (Sec. 5.2.2.3) for a two-outcome Bell inequality and that the Bell-CH inequality cannot be violated when any of the POVM elements considered are of full rank (Appendix B.3.1). Therefore, without loss of generality, we can restrict our attention to the following rank one projectors:

$$A_{s_a}^\pm = \frac{1}{2} (\mathbf{1}_2 \pm \hat{a}_{s_a} \cdot \vec{\sigma}), \quad (5.40)$$

where $\hat{a}_{s_a} \in \mathbb{R}^3$ for $s_a = 1, 2$ are unit vectors.

Next, we would like to optimize over Bob's measurements for this choice of Alice's measurement using Helstrom-like optimization [188] which has been discussed in Sec. 5.2.2.3. This allows us to obtain $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, A_{s_a}^{o_a})$ which can further be optimized to obtain $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$ using simple variational techniques. Substituting Eq. (5.37) and Eq. (5.40) into Eq. (5.13) and Eq. (5.15), and after some computation (Appendix B.3.2), it can be shown that for a general two-qubit state, Eq. (5.37),

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) = \max \left\{ 0, \frac{1}{2} \left(\sqrt{\varsigma_1^2 + \varsigma_2^2} - 1 \right) \right\}, \quad (5.41)$$

¹³We can easily obtain this particular representation from the coherence vector representation, Eq. (C.6a), by defining the rescaled basis matrices as $\sigma_0 = \mathbf{1}_2/\sqrt{2}$, $\sigma_1 = \sigma_x/\sqrt{2}$, $\sigma_2 = \sigma_y/\sqrt{2}$, $\sigma_3 = \sigma_z/\sqrt{2}$ and rescaling the various coefficients in Eq. (C.6a) as $\mathbf{r}_A \rightarrow \mathbf{r}_A/2$, $\mathbf{r}_B \rightarrow \mathbf{r}_B/2$, $R' \rightarrow T \equiv R'/2$.

¹⁴It should be clear that there is no unique way of writing the Bell operator derived from a given Bell inequality. The function that is of our interest, $\mathcal{S}_{\text{QM}}(\rho)$, however is in no way affected by this degeneracy.

where ς_1 and ς_2 are the two largest singular values of T . Since a Bell-CH violation for ρ occurs if and only if $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) > 0$, the necessary and sufficient condition for a two-qubit state ρ to violate the Bell-CH inequality, and hence the Bell-CHSH inequality [c.f. Eq. (5.36)] is

$$\varsigma_1^2 + \varsigma_2^2 > 1, \quad (5.42)$$

which is just the Horodecki criterion [30].¹⁵

5.3.3 I_{3322} -violation for a Class of Two-Qubit States

Next, we look at how the two algorithms can be used in tandem to determine if some two-qubit states violates the I_{3322} inequality introduced in Eq. (3.27). This Bell inequality is interesting in that there are quantum states that violate this new inequality but not the Bell-CH/Bell-CHSH inequality. The analogue of Horodecki's criterion for this inequality is thus very desirable.

To the best of our knowledge, such an analytic criterion is yet to be found. However, by combining the two algorithms presented above, we can often offer a definitive, yet nontrivial, conclusion about the compatibility of a quantum state with a locally causal description. For ease of reference, we will also reproduce the functional form of I_{3322} inequality as follows:

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_{3322})} = & p_{\text{AB}}^{o_a o_b}(1, 1) + p_{\text{AB}}^{o_a o_b}(1, 2) + p_{\text{AB}}^{o_a o_b}(1, 3) + p_{\text{AB}}^{o_a o_b}(2, 1) + p_{\text{AB}}^{o_a o_b}(2, 2) - p_{\text{AB}}^{o_a o_b}(2, 3) \\ & + p_{\text{AB}}^{o_a o_b}(3, 1) - p_{\text{AB}}^{o_a o_b}(3, 2) - p_{\text{A}}^{o_a}(1) - 2p_{\text{B}}^{o_b}(1) - p_{\text{B}}^{o_b}(2) \leq 0, \end{aligned} \quad (3.27)$$

where the outcomes o_a and o_b are labeled as “ \pm ”. Without loss of generality, we can restrict our attention to $o_a = o_b = “+”$. Then, from Eq. (5.3), Eq. (4.2) and Eq. (4.1a), it can be shown that the Bell operator corresponding to this Bell inequality reads:

$$\mathcal{B}_{I_{3322}} = A_1^+ \otimes (B_1^- - B_2^+ + B_3^-) - A_2^+ \otimes B_3^- - A_2^- \otimes (B_1^- + B_2^-) - A_3^+ \otimes B_2^- - A_3^- \otimes B_1^-. \quad (5.43)$$

For convenience, we will adopt the notation that $O_m^\pm \equiv A_m^\pm$ for $m = 1, 2, 3$ and $O_m^\pm \equiv B_{m-3}^\pm$ for $m = 4, 5, 6$. In these notations, the global optimization problem for this Bell inequality can be written as

$$\text{maximize} \quad \text{tr}(\rho \mathcal{B}_{I_{3322}}) \quad (5.44a)$$

$$\text{subject to} \quad (O_m^\pm)^2 = O_m \quad (5.44b)$$

for $m = 1, 2, \dots, 6$, which is a QCQP. The lowest order relaxation to this problem can thus be obtained by following similar procedures as that described in Sec. 5.2.3.

To obtain a *state-dependent* upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ for this inequality, we have to impose the analogue of Eq. (5.31b), i.e.,

$$z_m^\pm = \text{tr}(O_m^\pm) = 0, 1, \dots, d, \quad (5.45)$$

for each of the POVM elements. For small d , numerical upper bounds on $\mathcal{S}_{\text{QM}}(\rho)$ can then be solved for using SOSTOOLS. As an example, let's now look at how this upper bound,

¹⁵It is worth noting that yet another alternative derivation of Eq. (5.42) has also been given in Ref. [195].

together with the LB algorithm, has enabled us to determine if a class of mixed two-qubit states violates the I_{3322} inequality.

The mixed two-qubit state

$$\rho_{\text{CG}}(p) = p |\Psi_{2:1}\rangle\langle\Psi_{2:1}| + (1-p) |0\rangle_{\mathcal{A}\mathcal{A}}\langle 0| \otimes |1\rangle_{\mathcal{B}\mathcal{B}}\langle 1|, \quad 0 \leq p \leq 1, \quad (5.46)$$

can be understood as a mixture of the pure product state $|0\rangle_{\mathcal{A}}|1\rangle_{\mathcal{B}}$ and the non-maximally entangled two-qubit state $|\Psi_{2:1}\rangle = \frac{1}{\sqrt{5}}(2|0\rangle_{\mathcal{A}}|0\rangle_{\mathcal{B}} + |1\rangle_{\mathcal{A}}|1\rangle_{\mathcal{B}})$. As can be easily verified using the PPT criterion [161, 162], this state is entangled for $0 < p \leq 1$. In particular, the mixture with $p = 0.85$ was first presented in Ref. [60] as an example of a two-qubit state that violates the I_{3322} inequality but not the Bell-CH/ Bell-CHSH inequality.

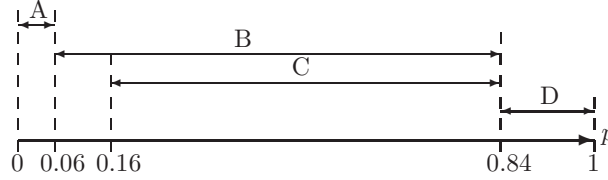


FIGURE 5.1: Domains of p where the compatibility between a locally causal description and quantum mechanical prediction given by $\rho_{\text{CG}}(p)$ was studied via the LB and UB algorithms in conjunction with the I_{3322} inequality. From right to left are respectively the domain of p whereby $\rho_{\text{CG}}(p)$ is: (D) found to violate the I_{3322} inequality; (C) found to give a lowest order upper bound that is compatible with the I_{3322} inequality; (B) found to give a higher order upper bound that is compatible with the I_{3322} inequality; (A) not known if it violates the I_{3322} inequality.

Given the above observation, a natural question that one can ask is, at what values of p does $\rho_{\text{CG}}(p)$ violate the I_{3322} inequality? Using the LB algorithm, we have found that for¹⁶ $p \gtrsim 0.83782$ (domain D in Figure 5.1), $\rho_{\text{CG}}(p)$ violates the I_{3322} inequality. As we have pointed out in Sec. 5.2.2, observables that lead to the observed level of I_{3322} -violation can be readily read off from the output of the SDP.

On the other hand, through the UB algorithm, we have also found that, with the lowest order relaxation, the states do not violate this 3-setting inequality for $0.16023 \lesssim p \lesssim 0.83625$ (domain C in Figure 5.1); with a higher order relaxation, this range expands to $0.06291 \lesssim p \lesssim 0.83782$ (domain B in Figure 5.1). Notice that at the presented accuracy, the upper bound of p where there can be no violation of the I_{3322} inequality now agrees with the lower bound of p where an I_{3322} violation was found.

The algorithms alone therefore leave a tiny gap at $0 < p \lesssim 0.06291$ (domain A in Figure 5.1) where we could not conclude if $\rho_{\text{CG}}(p)$ violates the I_{3322} inequality. Nevertheless, if we recall that the set of quantum states not violating a given Bell inequality is convex and that $\rho_{\text{CG}}(0)$, being a pure product state, cannot violate any Bell inequality, we can immediately conclude that $\rho_{\text{CG}}(p)$ with $0 \leq p \lesssim 0.83782$ cannot violate the I_{3322} inequality. As such, together with convexity arguments, the two algorithms allow us to fully characterize the state $\rho_{\text{CG}}(p)$ compatible with LHV Ts, when each observer is only allowed to perform three different dichotomic measurements.

¹⁶Throughout, we will use $p \gtrsim p'$ and $p \lesssim p'$ to denote p' as, respectively, a numerical (approximate) lower bound and upper bound for p .

5.3.4 Limitations of the UB algorithm

As can be seen in the above examples, the UB algorithm does not always provide a very good upper bound for $\mathcal{S}_{\text{QM}}(\rho)$. In fact, it has been observed that for pure product states, the algorithm with lowest order relaxation always returns a state-independent bound (the Tsirelson bound in the case of Bell-CHSH inequality). As such, for mixed states that can be decomposed as a high-weight mixture of pure product state and some other entangled state, the upper bound given by UB is typically bad. To illustrate this, let us consider the 1-parameter family of PPT bound entangled state $\rho_{\text{H}}(p)$ [160, 163], Eq. (4.21), and recall from Sec. 4.3.3.1 that a bipartite PPT entangled state cannot violate the Bell-CH or the Bell-CHSH inequality [165].

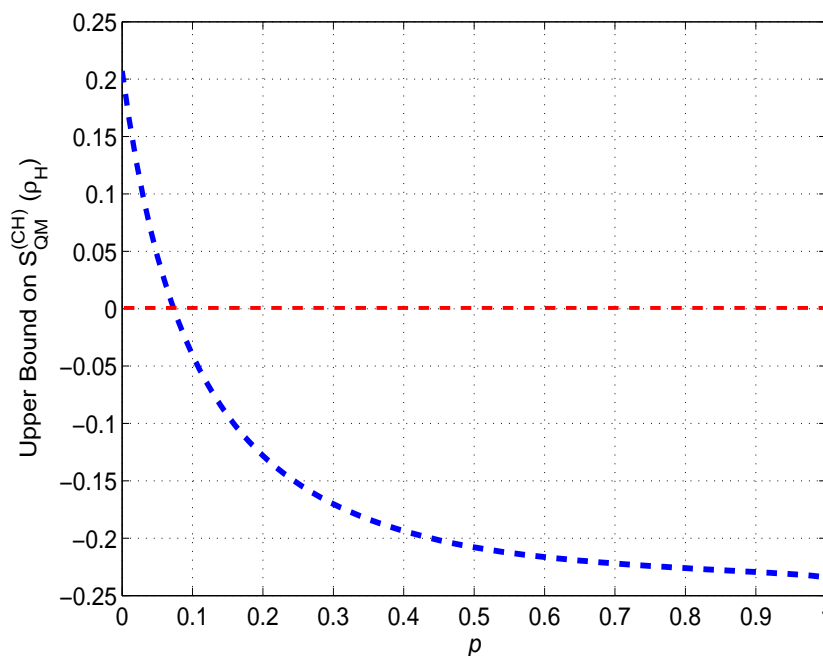


FIGURE 5.2: Numerical upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{\text{H}})$ obtained from the UB algorithm using lowest order relaxation and Eq. (5.45). The dashed horizontal line is the threshold above which no locally causal description is possible.

However, when tested with the UB algorithm using the lowest order relaxation, it turned out that some of these upper bounds are actually above the threshold of Bell-CH violation (see Figure 5.2). In fact, the upper bound obtained for the pure product state, $\rho_{\text{H}}(0) = |\Psi_p\rangle\langle\Psi_p|$ is actually the maximal achievable Bell-CH violation given by a quantum state [143]. Nonetheless, as with the example presented in Sec. 5.3.3, we can exclude the possibility of $\rho_{\text{H}}(p)$ violating the Bell-CH inequality by combining the upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{\text{H}})$ and the convexity of NBIV states.

5.4 Conclusion

In this chapter, we have looked specifically into the problem of determining if a given (entangled) quantum state is Bell-inequality-violating (BIV) for some fixed but arbitrary Bell inequality. For that purpose, we have presented two algorithms which can be used to determine, respectively, a lower bound (LB) and an upper bound (UB) on the maximal expectation value of a Bell operator for a given quantum state, i.e., $\mathcal{S}_{\text{QM}}(\rho)$.

In particular, we have demonstrated how one can make use of the upper bound to derive a necessary condition for two-qudit states with vanishing coherence vectors to violate the Bell-CHSH inequality. When $d = 2$, we have also illustrated how the LB algorithm can be used to rederive Horodecki's criterion for two-qubit states. For more complicated Bell inequalities where analytic treatment seems formidable, we have demonstrated how one can make use of the two algorithms in tandem to determine, numerically, if the quantum mechanical prediction is compatible with a locally causal description. In Chapter 6, we will also see how these algorithms have been applied to the search of maximal-Bell-inequality-violation in the context of collective measurements without postselection.

As with many other numerical optimization algorithms, the LB algorithm can only guarantee the convergence to a local maximum of $\mathcal{S}_{\text{QM}}(\rho, \{O_m\})$. The UB algorithm, on the other hand, provides an (often loose) upper bound on $\mathcal{S}_{\text{QM}}(\rho)$. In the event that these bounds agree (up to reasonable numerical precision), we know that optimization of the corresponding Bell operator using LB has been achieved. This ideal scenario, however, is not as common as we would like it to be. In particular, the UB algorithm with lowest order relaxation has been observed to give rather bad bounds for states with a high-weight mixture of pure product states (although we can often rule out the possibility of a violation in this situation by convexity arguments as in Sec. 5.3.3 and Sec. 5.3.4). A possibility to improve these bounds, as suggested by the work of Nie *et al.* [196], is to incorporate the *Karush-Kuhn-Tucker* optimality condition as an additional constraint to the problem. We have done some preliminary studies on this but have not so far found any improvement in the bounds obtained but this deserves further study.

As of now, we have only implemented the UB algorithm to determine upper bounds on $\mathcal{S}_{\text{QM}}(\rho)$ for dichotomic Bell inequalities. For Bell inequalities with more outcomes, the local Hermitian observables are generally also subjected to constraints in the form of a LMI. Although the UB algorithm can still be implemented for these Bell inequalities by first mapping the LMI to a series of polynomial inequalities, this approach seems blatantly inefficient. Future work to remedy this difficulty is certainly desirable.

Finally, despite the numerical and analytic evidence at hand, it is still unclear why the lowest order relaxation to the global optimization problem, as described in Sec. 5.2.3.2, seems always gives rise to a bound that is state-independent and how generally this is true. Some further investigation on this may be useful, particularly to determine whether the lowest order relaxation is always state-independent even for inequalities that are not correlation inequalities. If so this could complement the methods of Refs. [17, 56, 58] for finding state-independent bounds on Bell inequalities. In fact, recently, very similar techniques were found to give provably state-independent bounds on maximal Bell inequality violation [17, 197].

6

Bell-Inequality Violations by Quantum States

In this chapter, we will make use of the toolkits developed in Chapter 5 to analyze the extent to which specific quantum states can violate a given Bell inequality. Geometrically, the degree of violation of a given facet-inducing Bell inequality $I^{(k)}$ provides a measure of the distance of the Bell-inequality-violating quantum correlation from the boundary of the convex set of classical correlations corresponding to $I^{(k)}$. We will consider this problem both in the typical scenario where a quantum system is measured one copy at a time, and the other scenario where multiple copies of the same quantum system are measured collectively.

6.1 Introduction

Pioneering investigation on the extent to which a given quantum state can violate a given Bell inequality can be traced back to as early as 1980s. At that time, Mermin and Garg [102, 110, 111, 112] were mainly interested to know if this nonclassical feature displayed by two entangled spin- j quantum systems could survive in the “classical limit” of $j \rightarrow \infty$. Their initial attempt [110] seemed to have suggested that this nonclassical feature does indeed diminish with increasing quantum numbers, in agreement with the mentality that the classical world arises in the $j \rightarrow \infty$ limit. That this observation is an artefact of their analysis was almost immediately confirmed by their follow up work [102, 111], in which they showed that the spin- j singlet state for any j could indeed contradict predictions given by any LHVT.

A quantitative study of the strength of Bell-CHSH-violation for arbitrary spin- j singlet states was nonetheless not available until Peres revisited the problem almost a decade later [198]. The measurements that Peres considered in Ref. [198] are, however, not optimal and only lead to a Bell-CHSH-violation of 2.481 in the asymptotic limit of $j \rightarrow \infty$. This result was soon strengthened by Gisin and Peres [36], who showed that for the spin- j singlet state, i.e., the $(2j + 1)$ -dimensional maximally entangled state $|\Phi_{2j+1}^+\rangle$, the corresponding Bell-CHSH-violation is just $2\sqrt{2}$ (the Tsirelson bound) when j is a half integer, and tends

towards the same bound as $j \rightarrow \infty$ if j is an integer.

The strength of a Bell inequality violation is also relevant from an experimental point of view. Given that in a realistic experimental scenario, pure entangled states are hard, if not impossible, to prepare, a natural question that follows is the robustness of nonclassical correlations against the mixture of noise. How is the robustness of nonclassical correlations against noise related to the strength of violation? Crudely speaking, in the presence of noise, the strength of violation decreases, therefore the stronger an entangled state violates a given Bell inequality, the more robust are the corresponding nonclassical correlations against the mixture of noise. Along this line of investigation, Kaszlikowski and coauthors [199] made an interesting discovery that, as opposed to the mentality of $j \rightarrow \infty$ being the classical limit, the inconsistency between LHV and quantum mechanical prediction for $|\Phi_{2j+1}^+\rangle$ actually gets more robust against the mixture of noise as j increases.¹

Indeed, using the d -outcome CGLMP inequality that they derived, Collins *et al.* [121] showed that as d , the dimension of the local Hilbert space increases, the maximal violation found for $|\Phi_d^+\rangle$ against this class of inequalities also increases (see also Ref. [122]). This finding is, of course, consistent with the above intuition, and the discovery presented in Ref. [199] that as d increases, the nonclassical correlations derived from $|\Phi_d^+\rangle$ are more robust against the mixture of (white) noise. In this regard, it is also worth noting that, somewhat surprisingly, for a given d , $|\Phi_d^+\rangle$ is not the quantum state whose Bell inequality violation is most robust against the mixture of noise [16, 200, 201].

On the other hand, experiments to test Bell inequalities usually involve making many measurements on individual copies of the quantum system with the system being prepared in the same way for each measurement. In this chapter, we will also consider a somewhat different scenario and ask if *quantum nonlocality*² can be enhanced by making joint local measurements on multiple copies of the entangled state. We will use the maximal Bell inequality violation of a quantum state ρ as our measure of nonlocality. Our interest is to determine if $\rho^{\otimes N}$, when compared with ρ , can give rise to a higher Bell inequality violation for some $N > 1$.

A very similar problem was introduced by Peres [40] who considered Bell inequality violations under collective measurements but allowed the experimenters to make an auxiliary measurement on their systems and postselect on both getting a specific outcome of their measurement. Numerically, Peres showed that with collective measurements and postselection, a large class of two-qubit states give rise to better Bell inequality violation. However, note that the postselection in Peres' scheme is stronger than that in realistic Bell inequality experiments where detector inefficiencies require a postselection on events where both detectors fired. In such a case the failure probability is independent of the quantum state.

As with Peres' examples, existing results in the literature on nonlocality enhancement always involve some kind of postselection, it is thus of interest to investigate the power of collective measurements, without postselection, in terms of increasing Bell inequality violation. Indeed, it is one of the main purposes of this chapter to show that postselection

¹The noise is modeled by the incoherent mixture of the state in question with a maximally mixed state (see, for example, the discussion on $\rho_{W_d}(p)$ and $\rho_{I_d}(p)$ in Sec. 4.3.2.1 and Sec. 4.3.2.2).

²The term “quantum nonlocality” is used here merely as a widely, but not universally accepted synonym for the violation of a Bell inequality (see e.g. Ref. [202] and Ref. [149] for opposing views).

is not necessary to improve Bell inequality violation. In order to find such examples for mixed states we have resorted to various numerical approaches that are described in Sec. 5.2 to provide upper (UB) and lower bounds (LB) on the optimal violation of a given Bell inequality by a given quantum state. Unless otherwise stated, Bell inequality violations presented hereafter refer to the best violation that we could find either analytically, or numerically using the LB algorithm. For ease of reference, upper bounds obtained via UB are marked where they appear with † . In the event that a violation presented is known to be maximal (such as those computable using the Horodecki's criterion [30]), an $*$ will be attached.

This chapter is organized as follows. In Sec. 6.2.1, we present a measurement scheme which we will use to determine the Bell-CH inequality violation for any bipartite pure state. These measurements led to the largest violation that we were able to find and may even be maximal. This is then followed by a review of what is known about the best I_{22nn} -violation for some two-qudit states in Sec. 6.2.2. Then, in Sec. 6.3.1, we show that for bipartite pure entangled states, collective measurement can lead to a greater violation of the Bell-CH inequality. The corresponding scenario for mixed entangled states is analyzed in Sec. 6.3.2. We then conclude with a summary of results and some future avenues of research.

6.2 Single Copy Bell Inequality Violation

6.2.1 Bell-CH-violation for Pure Two-Qudit States

In this section, we present a measurement scheme which gives rise to the largest Bell-CH inequality violation that we have found for arbitrary pure two-qudit states. We find using this inequality for probabilities rather than correlations to be convenient for our purposes. From Eq. (5.36), we know that if the conjectured measurement scheme is optimal for the Bell-CH inequality, it will also give rise to the maximal Bell-CHSH inequality violation for any pure two-qudit state.

For ease of reference, let us again reproduce the functional form of Bell-CH inequality here:

$$\mathcal{S}_{\text{LHV}}^{(\text{CH})} = p_{\mathcal{AB}}^{o_a o_b}(1, 1) + p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{AB}}^{o_a o_b}(2, 2) - p_{\mathcal{A}}^{o_a}(1) - p_{\mathcal{B}}^{o_b}(1) \leq 0, \quad (3.23)$$

where in quantum mechanics, the relevant joint and marginal probabilities are calculated according to Eq. (4.2). Without loss of generality, in the following discussion, we will focus on the above inequality with $o_a = o_b = "+"$.

The maximal Bell inequality violation for a quantum state is invariant under a local unitary transformation. As such, the maximal Bell inequality violation for any bipartite pure quantum state is identical to its maximal violation when written in the Schmidt basis [203, 204]. In this basis, an arbitrary pure two-qudit state, i.e., $|\Phi_d\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ takes the form

$$|\Phi_d\rangle = \sum_{i=1}^d c_i |i\rangle_{\mathcal{A}} |i\rangle_{\mathcal{B}}, \quad (6.1)$$

where $\{|i\rangle_{\mathcal{A}}\}$ and $\{|i\rangle_{\mathcal{B}}\}$ are local orthonormal bases of subsystem possessed by observer \mathcal{A} and \mathcal{B} respectively, and $\{c_i\}_{i=1}^d$ are the Schmidt coefficients of $|\Phi_d\rangle$. Without loss of

generality, we may also assume that $c_1 \geq c_2 \geq \dots \geq c_d > 0$. Then $|\Phi_d\rangle$ is entangled if and only if $d > 1$. Now, let us consider the following measurement settings for Alice, which were first adopted in Ref. [36],³

$$\begin{aligned} A_1^\pm &= \frac{1}{2} [\mathbf{1}_d \pm Z], & A_2^\pm &= \frac{1}{2} [\mathbf{1}_d \pm X], \\ Z &\equiv \bigoplus_{i=1}^{\lfloor d/2 \rfloor} \sigma_z + \Xi, & X &\equiv \bigoplus_{i=1}^{\lfloor d/2 \rfloor} \sigma_x + \Xi, \\ [\Xi]_{ij} &= 0 \quad \forall \quad i, j \neq d, & [\Xi]_{dd} &= d \pmod{2}, \end{aligned} \quad (6.2)$$

where σ_x and σ_z are respectively the Pauli x and z matrices introduced in Eq. (2.6).

Notice, however, that the $\{B_{s_b}^\pm\}_{s_b=1}^2$ given in Ref. [36] are not optimal for a general pure two-qudit state. In fact, as we have seen in Sec. 5.2.2.3, given the measurements for Alice in Eq. (6.2), the optimization of Bob's measurement settings can be carried out explicitly. Using the resulting analytic expression for Bob's optimal POVM (Appendix B.4.1), the optimal expectation value of the Bell-CH operator, Eq. (5.39), for $|\Phi_d\rangle$ can be computed and we find

$$\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d\rangle} = \frac{1}{2} \sum_{n=1}^{\lfloor d/2 \rfloor} \sqrt{(c_{2n-1}^2 + c_{2n}^2)^2 + 4c_{2n}^2 c_{2n-1}^2} + \frac{\xi}{2} c_d^2 - \frac{1}{2}, \quad (6.3)$$

where $\xi \equiv d \pmod{2}$. From here, it is easy to see that for any entangled $|\Phi_d\rangle$, i.e., $d > 1$,

$$\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d\rangle} > \frac{1}{2} \sum_{n=1}^{\lfloor d/2 \rfloor} \sqrt{(c_{2n-1}^2 + c_{2n}^2)^2} + \frac{\xi}{2} c_d^2 - \frac{1}{2} = 0, \quad (6.4)$$

where we have made use of the normalization condition $\sum_{i=1}^d c_i^2 = 1$. Therefore, as was first shown by Gisin and Peres [36], a pure two-qudit state violate the Bell-CH, or equivalently the Bell-CHSH inequality if and only if it is entangled.

Effectively, the measurement scheme presented above corresponds to first ordering each party's local basis vectors $\{|i\rangle\}_{i=1}^d$ according to their Schmidt coefficients, and grouping them pairwise in descending order from the Schmidt vector with the largest Schmidt coefficient. Physically, this can be achieved by Alice and Bob each performing an appropriate local unitary transformation. Each of their Hilbert spaces can then be represented as a direct sum of 2-dimensional subspaces, which can be regarded as a one-qubit space, plus a 1-dimensional subspace if d is odd. The final step of the measurement consists of performing the optimal measurement ([30], see also Appendix B.3.2) in each of these two-qubit spaces as if the other spaces did not exist.

From here, it is easy to see that if we have a d -dimensional maximally entangled state

³Here, as well as Eq. (B.19) and Eq. (B.20), we will adopt the convention that when d is odd, the end product of the direct sum is appended with zero entries to make the dimension of the resulting matrix $d \times d$.

$|\Phi_d^+\rangle$, Eq. (4.17), then Eq. (6.3) gives⁴

$$\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d^+\rangle} = \begin{cases} \frac{1}{\sqrt{2}} - \frac{1}{2}^* : d \text{ even} \\ \frac{\sqrt{2}(d-1)+1}{2d} - \frac{1}{2} : d \text{ odd} \end{cases}. \quad (6.5)$$

With this measurement scheme, the Bell-CH inequality violation for a maximally entangled state with even d is thus the maximum allowed by Tsirelson's bound [143] whereas that of maximally entangled state with odd d is not. This may seem surprising at first glance, but as was pointed out by Popescu and Rohrlich in Ref. [180], the Tsirelson bound can *never* be attained by any $|\Phi_d\rangle$ with odd d .

How good is the measurement scheme given by Eq. (6.2) and Eq. (B.21)? It is constructed so that for pure two-qubit states, i.e. when $d = 2$, Eq. (6.3) gives the same violation found in Refs. [35, 36], and is the maximal violation determined by Horodecki *et al.* [30] (Appendix B.3.2). The measurement given by Eq. (6.2) is hence optimal for any two-qubit state $|\Phi_2\rangle$. Moreover, for the 3-dimensional isotropic state $\rho_{\text{I}_3}(p)$, c.f. Eq. (4.16),

$$\rho_{\text{I}_3}(p) = p |\Phi_3^+\rangle\langle\Phi_3^+| + (1-p) \frac{\mathbf{1}_3 \otimes \mathbf{1}_3}{9}, \quad (6.6)$$

the measurement scheme given by Eq. (6.2) and Eq. (B.21) gives rise to

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{\text{I}_3}) = \max \left\{ \left(\frac{1+3\sqrt{2}}{9} \right) p - \frac{4}{9}, 0 \right\}, \quad (6.7)$$

which is exactly the maximum Bell-CH violation of $\rho_{\text{I}_3}(p)$ as determined by Ito *et al.* [14]. In other words, the measurement operators given by Eq. (6.2) and Eq. (B.21) are also optimal for $|\Phi_3^+\rangle$ and its mixture with the maximally mixed state.

In general, for higher dimensional quantum systems, we have looked at randomly generated pure two-qudit states ($d = 3, \dots, 10$) with their (unnormalized) Schmidt coefficients uniformly chosen at random from the interval $(0, 1)$. For all the 20,000 states generated for each d , we found that with Eq. (6.2) as the initial measurement setting, the (iterative) LB algorithm never gives a $\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d\rangle}$ that is different from Eq. (6.3) by more than 10^{-15} , thus indicating that Eq. (6.3) is, at least, a local maximum of the optimization problem.

Furthermore, for another 8,000 randomly generated pure two-qudit states, 1,000 each for $d = 3, \dots, 10$, an extensive numerical search using more than 4.6×10^6 random initial measurement guesses have not led to a single instance where $\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d\rangle}$ is higher than that given in Eq. (6.3)⁵. These numerical results suggest that the measurement scheme given by Eq. (6.2) and Eq. (B.21) may be the optimal measurement that maximizes the Bell-CH inequality violation for arbitrary pure two-qudit states.

⁴Although Bob's measurements $\{B_{s_b}^\pm\}_{s_b=1}^2$ given in Ref. [36] are generally not optimal when Alice's measurements are given by Eq. (6.2), the measurement settings given in Ref. [36] do give rise to the same $\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d^+\rangle}$ as we have got here for maximally entangled state.

⁵It is worth noting that among the 1,000 random pure states generated for each d , there are always some whose best Bell-CH inequality violation found differs from Eq. (6.3) by no more than 10^{-10} .

6.2.2 CGLMP and I_{22nn} -violation for Some Two-Qudit States

Apart from the Bell-CH/ Bell-CHSH inequalities, the other class of bipartite Bell inequalities whose quantum violations are most well-studied in the literature is probably the CGLMP inequality, Eq. (3.36), which is equivalent to the I_{22nn} inequality, Eq. (3.35). For any quantum state ρ , its violations of these two inequalities are shown in Appendix B.1.1 to be related linearly as follows:

$$\text{tr}(\rho \mathcal{B}_{I_n}) = \frac{2n}{n-1} \text{tr}(\rho \mathcal{B}_{I_{22nn}}) + 2, \quad (6.8)$$

where \mathcal{B}_{I_n} is the Bell operator derived from the n -outcome CGLMP inequality, Eq. (3.36). In Eq. (6.8), $\mathcal{B}_{I_{22nn}}$ is the Bell operator associated with the I_{22nn} inequality, which can be written explicitly as

$$\begin{aligned} \mathcal{B}_{I_{22nn}} = & \sum_{o_a=1}^{n-1} \sum_{o_b=1}^{n-o_a} A_1^{o_a} \otimes B_1^{o_b} + \sum_{o_a=1}^{n-1} \sum_{o_b=n-o_a}^{n-1} \left(A_1^{o_a} \otimes B_2^{o_b} + A_2^{o_a} \otimes B_1^{o_b} - A_2^{o_a} \otimes B_2^{o_b} \right) \\ & - \sum_{o_a=1}^{n-1} A_1^{o_a} \otimes \mathbb{1}_{d_B} - \sum_{o_b=1}^{n-1} \mathbb{1}_{d_A} \otimes B_1^{o_b}, \end{aligned} \quad (6.9)$$

where d_A and d_B are, respectively, the dimension of Alice's and Bob's Hilbert spaces.

In this section, we will give a brief review of the best CGLMP-violation and hence — via Eq. (6.8) — the best I_{22nn} -violation known for the isotropic state $\rho_{I_d}(p)$,

$$\rho_{I_d}(p) = p |\Phi_d^+\rangle\langle\Phi_d^+| + (1-p) \frac{\mathbb{1}_d \otimes \mathbb{1}_d}{d^2}, \quad (4.16)$$

where p is the weight of the d -dimensional maximally entangled state $|\Phi_d^+\rangle$ in the mixture. In what follows, we shall thus be contented with the scenario where $d_A = d_B = n = d$. Interestingly, it turned out that the best known I_{22nn} -violation for $\rho_{I_d}(p)$ is achieved with rank-one projective measurements. By linearity of expectation value, it therefore suffices to determine the maximal I_{22dd} -violation for $|\Phi_d^+\rangle$; the best I_{22dd} -violation for $\rho_{I_d}(p)$ will follow immediately. These best known violations will come in handy when we need to compare the best I_{22dd} -violation that we have found against what is known in the literature.

Now, let us recall the best known I_{22dd} -violation for $|\Phi_d^+\rangle$. From the pioneering result of Collins *et al.* ([121], see also Ref. [122]), it follows that with rank-one projective measurements, the d -dimensional maximally entangled state $|\Phi_d^+\rangle$ can violate the I_{22dd} inequality by as much as

$$\langle \mathcal{B}_{I_{22dd}} \rangle_{|\Phi_d^+\rangle} = \frac{d-1}{2d} \left[4d \sum_{k=0}^{\lfloor \frac{d}{2} \rfloor - 1} (q_k - q_{-(k+1)}) - 2 \right], \quad (6.10)$$

where $q_k \equiv \frac{1}{2d^3 \sin^2[\pi(k+\frac{1}{4})d]}$. In particular, in the asymptotic limit of $d \rightarrow \infty$, this best I_{22dd} -violation by $|\Phi_d^+\rangle$ converges to

$$\lim_{d \rightarrow \infty} \langle \mathcal{B}_{I_{22dd}} \rangle_{|\Phi_d^+\rangle} = \frac{1}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(k+1/4)^2} - \frac{1}{(k+3/4)^2} = \frac{16}{\pi^2} \times \text{Catalan} - 1 \approx 0.484 \, 91 \quad (6.11)$$

where $\text{Catalan} \approx 0.915\,97$ is the Catalan constant. Explicit values for some of these best known violations can be found in column 4 of Table 6.1. From column 2 and 3 of the same table, it can also be seen that the best known violation of this inequality is apparently not attained by the maximally entangled state $|\Phi_d^+\rangle$ — an interesting phenomenon that was first discovered by Acín *et al.* [16].

Table 6.1: Best known CGLMP-violation and I_{22dd} -violation for the maximally entangled two-qudit state $|\Phi_d^+\rangle$. The first column of the table gives the dimension of the local subsystem d . The second column gives the largest possible quantum violation of the CGLMP inequality for $d \leq 8$, first obtained in Ref. [16], and subsequently verified in Ref. [17]; these maximal violations also set an upper bound on the maximal violation attainable by $|\Phi_d^+\rangle$ for each d . The third column of the table gives the best known d -outcome CGLMP-violation for $|\Phi_d^+\rangle$ whereas the fourth column gives the corresponding best known I_{22dd} -violation obtained from Eq. (6.8). Also included in the fifth column of the table is the threshold weight p_d below which no violation of either inequality by isotropic state $\rho_{I_d}(p)$ is known.

d	$\mathcal{S}_{\text{QM}}^{(\text{CGLMP})}(\rho)$	$\langle \mathcal{B}_{\text{CGLMP}} \rangle_{ \Phi_d^+\rangle}$	$\langle \mathcal{B}_{I_{22dd}} \rangle_{ \Phi_d^+\rangle}$	p_d
2	2.8284	2.8284	0.20711	0.70711
3	2.9149	2.8729	0.29098	0.69615
4	2.9727	2.8962	0.33609	0.69055
5	3.0157	2.9105	0.36422	0.68716
8	3.1013	2.9324	0.40793	0.68203
10	-	2.9398	0.42291	0.68032
100	-	2.9668	0.47856	0.67413
1000	-	2.9695	0.48427	0.67351
∞	-	2.9698	0.48491	0.67349

Now, it is not difficult to see from Eq. (6.9) that when restricted to rank-one projective measurements, the expectation value of $\mathcal{B}_{I_{22nn}}$ with respect to the $d \times d$ -dimensional maximally mixed state $\rho_{d \times d}$ reads:

$$\text{tr}(\rho_{d \times d} \mathcal{B}_{I_{22dd}}) = -1 + \frac{1}{d}. \quad (6.12)$$

Therefore, from the linearity of expectation value and Eq. (6.10), it follows that the best known I_{22dd} -violation for the isotropic states is:

$$\text{tr}[\rho_{I_d}(p) \mathcal{B}_{I_{22dd}}] = p \times \frac{d-1}{2d} \left[4d \sum_{k=0}^{\lfloor \frac{d}{2} \rfloor - 1} (q_k - q_{-(k+1)}) - 2 \right] + (1-p) \left(-1 + \frac{1}{d} \right), \quad (6.13)$$

On the other hand, given that this best known violation increases linearly with p , it is also easy to see that there exists a threshold weight $p = p_d$ (sometimes called the visibility parameter) below which $\rho_{I_d}(p)$ is not known to violate the I_{22dd} inequality. Explicit values for some of these threshold weights can be found in column 6 of Table 6.1. In principle, it is of course possible that $\rho_{I_d}(p)$ with $1/(d+1) \leq p < p_d$ violates I_{22dd} and/or other Bell

inequalities for $\mathcal{P}_{2;2}^{d;d}$ with some other choice of measurements. However, preceding results due to Kaslikowski *et al.* [199] suggest that p_d could very well be the threshold p below which $\rho_{I_d}(p)$ does not violate any Bell inequalities for $\mathcal{P}_{2;2}^{d;d}$ (see also Refs. [200, 201, 205] in this regard). In other words, Eq. (6.13) may very well give the maximal I_{22dd} -violation for the isotropic states.

6.3 Better Bell-inequality Violation by Collective Measurements

6.3.1 Multiple Copies of Pure States

Let us now look into the problem of whether stronger nonclassical correlations can be derived by performing collective measurements on $N > 1$ copies of an entangled quantum state⁶. As our first example of nonlocality enhancement, consider again those pure maximally entangled two-qudit states residing in Hilbert space with odd d . As remarked earlier, it is well-known that their maximal Bell-CH/ Bell-CHSH inequality violation cannot saturate Tsirelson's bound [180]. In fact, their best known Bell-CH inequality violation [36] is that given in Eq. (6.5). By combining N copies of these quantum states, it is readily seen that we effectively end up with another maximally entangled state of d^N -dimension. It then follows from Eq. (6.5) that their Bell-CH violation under collective measurements increases monotonically with the number of copies N (see also Table 6.2, column 3 and 7). In fact, it can be easily shown that this violation approaches asymptotically the Tsirelson bound [143] in the limit of large N . Therefore, if the maximal violation of these quantum states is given by Eq. (6.5), which is the case for $d = 3$ [14], collective measurements can already give better Bell-CH violation with $N = 2$. Even if the maximal violation is not given by Eq. (6.5), it can be seen (by comparing the upper bound of the single-copy violation from the UB algorithm and the lower bound of the N -copy violation) from Table 6.2 that for $d = 5$, a Bell-CH violation better than the maximal single-copy violation can always be obtained when N is sufficiently large.

Such an enhancement is even more pronounced in the case of non-maximally entangled states. In particular, for N copies of a (non-maximally entangled) two-qubit state written in the Schmidt basis,

$$|\Phi_2\rangle^{\otimes N} = (\cos \phi |0\rangle_{\mathcal{A}} |0\rangle_{\mathcal{B}} + \sin \phi |1\rangle_{\mathcal{A}} |1\rangle_{\mathcal{B}})^{\otimes N}, \quad (6.14)$$

where $0 < \phi \leq \frac{\pi}{4}$ ⁷, the Bell-CH violation given by Eq. (6.3) is

$$\langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_2\rangle} = \frac{p}{\sqrt{2}} + \frac{1-p}{2} \sqrt{1 + \sin^2 2\phi} - \frac{1}{2}, \quad (6.15)$$

⁶Notice that the maximal Bell inequality violation for $N > M$ copies of a quantum system is never less than that involving only M copies. This follows from the fact that the maximal M -copy violation can always be recovered in the N -copy scenario by performing the M -copy-optimal-measurement on M of the N copies, while leaving the remaining $N - M$ copies untouched.

⁷For $\frac{\pi}{4} < \phi < \frac{\pi}{2}$, we just have to redefine ϕ as $\frac{\pi}{2} - \phi$ and all the subsequent results follow.

where

$$p = 1 - \frac{1}{2} \cos^{2(N-1)} \phi \sum_{m=0}^{N-1} \tan^{2m} \phi \left[1 - (-1)^{\frac{(N-1)!}{m!(N-1-m)!}} \right],$$

is the total probability of finding $|\Phi_2\rangle^{\otimes N}$ in one of the *perfectly correlated 2-dimensional subspaces* (i.e., a subspace with $c_{2n-1} = c_{2n}$) upon reordering of the Schmidt coefficients in descending order.

It is interesting to note that for these two-qubit states, their Bell-CH inequality violation for $N = 2k - 1$ copies, and $N = 2k$ copies are identical⁸ for all $k \geq 1$, as illustrated in column 2 of Table 6.2 and in Figure 6.1. This feature, however, does not seem to generalize to higher dimensional quantum states.

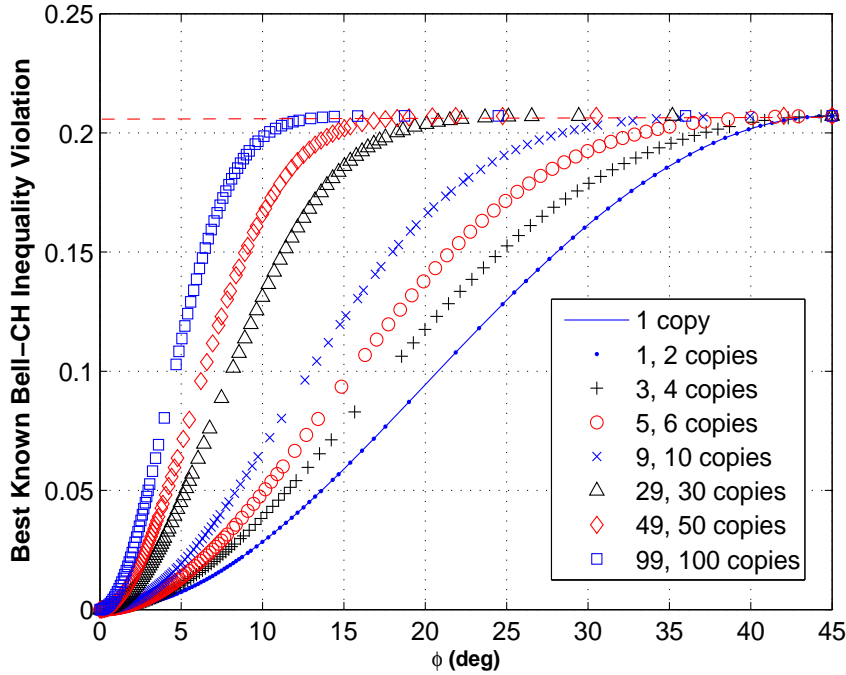


FIGURE 6.1: Best known Bell-CH inequality violation of pure two-qubit states obtained from Eq. (6.3), plotted as a function of ϕ , which gives a primitive measure of entanglement; $\phi = 0$ for bipartite pure product state and $\phi = 45^\circ$ for bipartite maximally entangled state. The curves from right to left represent increasing numbers of copies. The dotted horizontal line at $\frac{1}{\sqrt{2}} - \frac{1}{2}$ is the maximal possible violation of Bell-CH inequality; correlations allowed by locally causal theories have values less than or equal to zero. The solid line is the maximal Bell-CH inequality violation of $|\Phi_2\rangle$ determined using the Horodecki criterion, c.f. Appendix B.3.2.

Like the odd-dimensional maximally entangled state, the violation of the Bell-CH inequality for *any* pure two-qubit entangled states, as given by Eq. (6.3), increases asymptotically towards the Tsirelson bound [143] with the number of copies N , as can be seen in Figure 6.1.

⁸This can be rigorously shown using combinatoric arguments (private communication, Henry Haselgrove).

Table 6.2: Best known Bell-CH inequality violation for some bipartite pure entangled states, obtained from Eq. (6.2) and Eq. (B.21) with and without collective measurements. Also included below is the upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(|\Phi\rangle\langle\Phi|)$ obtained from the UB algorithm. Each of these upper bounds is marked with a † . The first column of the table gives the number of copies N involved in the measurements. Each quantum state is labeled by its non-zero Schmidt coefficients, which are separated by $:$ in the subscripts attached to the ket vectors; e.g., $|\Phi\rangle_{3:3:2:1}$ is the state with unnormalized Schmidt coefficients $\{c_i\}_{i=1}^4 = \{3, 3, 2, 1\}$. For each quantum state there is a box around the entry corresponding to the smallest N such that the lower bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(|\Phi\rangle\langle\Phi|)$ exceeds the single-copy upper bound (coming from the UB algorithm or otherwise). A violation that is known to be maximal is marked with a $*$.

N	$ \Phi_{2:1}\rangle$	$ \Phi_{1:1:1}\rangle$	$ \Phi_{3:2:1}\rangle$	$ \Phi_{4:3:2:1}\rangle$	$ \Phi_{3:3:2:1}\rangle$	$ \Phi_{1:1:1:1:1}\rangle$
	Lower	Bound				
1	0.14031*	0.13807*	0.16756	0.18431	0.19259	0.16569
2	0.14031	0.18409	0.18307	0.19624	0.20516	0.19882
3	0.16169	0.19944	0.19451	0.20275	0.20685	0.20545
4	0.16169	0.20455	0.19642	0.20388	0.20706	0.20678
5	0.17964	0.20625	0.20254	0.20596	0.20710	0.20704
10	0.19590	0.20710	0.20643	0.20704	0.20711	0.20711
	Upper	Bound				
1	0.14031*	0.13807*	0.19624 †	0.20711 †	0.20711 †	0.20569 †

Similarly, if we consider N copies of pure two-qutrit entangled states written in the Schmidt form,

$$|\Phi_3\rangle^{\otimes N} = (\cos \phi |0\rangle_{\mathcal{A}} |0\rangle_{\mathcal{B}} + \sin \phi \cos \theta |1\rangle_{\mathcal{A}} |1\rangle_{\mathcal{B}} + \sin \phi \sin \theta |2\rangle_{\mathcal{A}} |2\rangle_{\mathcal{B}})^{\otimes N}, \quad (6.16)$$

where $0 < \phi \leq \frac{\pi}{4}$, $0 < \theta \leq \frac{\pi}{4}$, it can be verified that their Bell-CH inequality violation, as given by Eq. (6.3), also increases steadily with the number of copies. Thus, if Eq. (6.3) gives the maximal Bell-CH violation for pure two-qutrit states, better Bell-inequality violation can also be attained by collective measurements using two copies of these quantum states. The explicit value of the violation can be found in column 3 and 4 of Table 6.2 for two specific two-qutrit states. As above, even if the maximal Bell-CH violation is not given by Eq. (6.3), collective measurements with Eq. (6.2) can definitely give a violation that is better than the maximal-single-copy ones as a result of the bound coming from the UB algorithm for a single copy (see Table 6.2). Corresponding examples for pure bipartite 4-dimensional and 5-dimensional quantum states can also be found in the table.

Some intuition for the way in which better Bell-CH inequality violation may be obtained with collective measurements and the measurement scheme given by Eq. (6.2) and Eq. (B.21) is that the reordering of subspaces prior to these measurements generally increases the total probability of finding 2-dimensional subspaces with $c_{2n} = c_{2n-1}$, while ensuring that the remaining 2-dimensional subspaces are at least as correlated as any of the corresponding single-copy 2-dimensional subspaces. The measurement then effectively projects onto each

of these subspaces (with Alice and Bob being guaranteed to obtain the same result) and then performs the optimal measurement on the resulting shared two-qubit state. Since the optimal measurements in each of these perfectly correlated 2-dimensional subspaces gives the maximal Bell-CH inequality violation, while the same measurements in the remaining 2-dimensional subspaces give as much violation as the single-copy violation, the multiple-copy violation is thus generally greater than that of a single copy.

As one may have noticed, our measurement protocol bears some resemblance with the entanglement concentration protocol developed by Bennett *et al.* [206]. In entanglement concentration, Alice and Bob make slightly different projections onto subspaces that are spanned by all those ket vectors sharing the same Schmidt coefficients thus obtaining a maximally entangled state in a bipartite system of some dimension. One can also obtain improved Bell inequality violations by adopting their protocol and first projecting Alice's Hilbert space into one of the perfectly correlated subspaces and performing the best known measurements for a Bell inequality violation in each of these (not necessary 2-dimensional) subspaces. We have compared the Bell-CH inequality violation of an arbitrary pure two-qubit state derived from each of these protocols and found that the violation obtained using our protocol always outperforms the one based on entanglement concentration. The difference, nevertheless, diminishes as $N \rightarrow \infty$. This observation provides another consistency check of the optimality of Eq. (6.3).

6.3.2 Multiple Copies of Mixed States

The impressive enhancement in a pure state Bell-CH inequality violation naturally leads us to ask if the same conclusion can be drawn for mixed entangled states. The possibility of obtaining better Bell inequality violation with collective measurements, however, does not seem to generalize to all entangled states.

Our first counterexample comes from the 2-dimensional Werner state, Eq. (4.12), which can seen as a mixture of the spin- $\frac{1}{2}$ singlet state and the maximally mixed state,

$$\rho_{w_2}(p) = p |\Psi^-\rangle\langle\Psi^-| (1-p) \frac{\mathbb{1}_2 \otimes \mathbb{1}_2}{4}, \quad (6.17)$$

where p is the weight of $|\Psi^-\rangle$ in the mixture. This state is entangled for $p > 1/3$ (c.f. Sec. 4.3.2.1) and from the Horodecki criterion (Appendix B.3.2) one can easily show that it violates the Bell-CH inequality if and only if [30]

$$p > p_w \equiv \frac{1}{\sqrt{2}} \simeq 0.707 \ 11 \quad (6.18)$$

Using the LB algorithm, we have searched for the maximal violation of $\rho_{w_2}(p)$ with $p > p_w$ for $N \leq 4$ copies but no increase in the maximal violation of Bell-CH inequality has ever been observed (see Figure 6.2). In fact, by using the UB algorithm, we find that for two copies of some Bell-CH violating Werner states, their maximal Bell-CH inequality violation are identical to the corresponding single-copy violation within a numerical precision of 10^{-12} . This strongly suggests that for some Werner states the maximal Bell-CH inequality violation does not depend on the number of copies N .

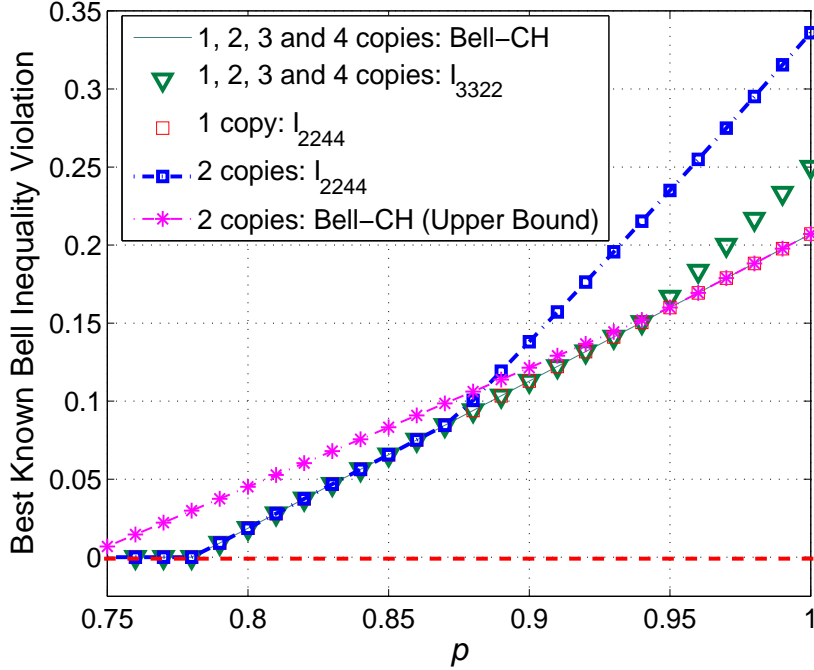


FIGURE 6.2: Best known expectation value of the Bell operator coming from the Bell-CH inequality [\mathcal{B}_{CH} , Eq. (5.39)], I_{3322} inequality [$\mathcal{B}_{I_{3322}}$, Eq. (5.43)], and the I_{2244} inequality [$\mathcal{B}_{I_{2244}}$, Eq. (6.9)] with respect to the 2-dimensional Werner states $\rho_{W_2}(p)$; p represents the weight of the spin- $\frac{1}{2}$ singlet state in the mixture. Note that the best I_{2244} -violation found for $|\Psi^-\rangle^{\otimes 2}$ agrees with the best known violation presented in Table 6.1. Also included is the upper bound on $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{W_2}^{\otimes 2})$ obtained from the UB algorithm.

There are, nevertheless, some two-qubit states whose maximal Bell-CH inequality violation for $N = 3$ is higher than the corresponding single-copy violation. In contrast to the pure state scenario, the set of mixed two-qubit states seems to be dominated by those whose 3-copy Bell-CH inequality violation is not enhanced. In fact, among 50,000 randomly generated Bell-CH violating two-qubit states⁹, only about 0.38% of them were found to have their 3-copy Bell-CH inequality violation greater than their maximal single-copy violation. Moreover, as can be seen in Figure 6.3, they are all clustered at regions with relatively low linear entropy.

As with the pure state scenario, an enhancement of nonclassical correlations in the Bell-CH setting seems to be more prevalent in higher dimensional quantum systems. In particular, for all the 3-dimensional isotropic states [Eq. (6.6)] that violate the Bell-CH inequality, numerical results obtained from the LB algorithm suggest that the maximal violation increases steadily with the number of copies. The results are summarized in Figure 6.4.

Yet another question that one can ask is how much does the enhancement of nonclassical correlations depend on the choice of Bell inequality. To address this question, we have also

⁹We follow the algorithm presented in Ref. [207] to generate random two-qubit states. In particular, the eigenvalues $\{\lambda_i\}_{i=1}^4$ of the quantum states were chosen from a uniform distribution on the 4-simplex defined by $\sum_i \lambda_i = 1$.

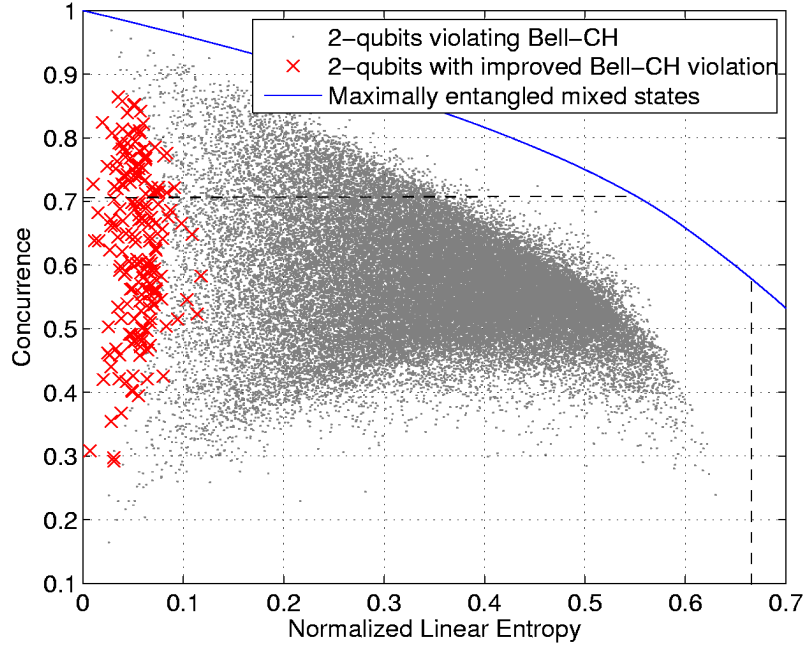


FIGURE 6.3: Distribution of two-qubit states sampled for better Bell-CH violation by collective measurements. The maximally entangled mixed states (MEMS), which demarcate the boundary of the set of density matrices on this concurrence-entropy plane [9, 10], are represented by the solid line. Note that as a result of the chosen distribution over mixed states this region is not well sampled. The region bounded by the solid line and the horizontal dashed line (with concurrence equal to $1/\sqrt{2}$) only contain two-qubit states that *violate* the Bell-CH inequality [11]; the region bounded by the solid line and the vertical dashed line (with normalized linear entropy equal to $2/3$) only contain states that *do not violate* the Bell-CH inequality [11, 12, 13]. Two-qubit states found to give better 3-copy Bell-CH violation are marked with red crosses.

studied the enhancement of nonclassical correlations with respect to other Bell inequalities for probabilities, in particular the I_{3322} inequality given in Eq. (3.27), the I_{2233} inequality given in Eq. (3.32) and the I_{2244} inequality given in Eq. (3.34). For these Bell inequalities, we find that the possibility of enhancing nonclassical correlations does seem to depend on both the number of alternative settings and the number of possible outcomes involved in a Bell experiment. The dependence on the number of outcomes is particularly prominent in the case of 2-dimensional Werner states, where a large range of I_{2244} -inequality-violating $\rho_{W_2}(p)$ seem to achieve a higher two-copy violation, even though their maximal Bell-CH inequality violation apparently remains unchanged up to $N = 4$ (Figure 6.2).

The dependence on the number of alternative settings can be seen in the best known violation of $\rho_{I_3}(p)$ with respect to the Bell-CH inequality and the I_{3322} inequality (Figure 6.4). In particular, when the number of alternative settings is increased from 2 (in the case of Bell-CH inequality) to 3 (in the case of I_{3322} inequality), the range of states whereby collective measurements were found to improve the Bell inequality violation is drastically *reduced*.

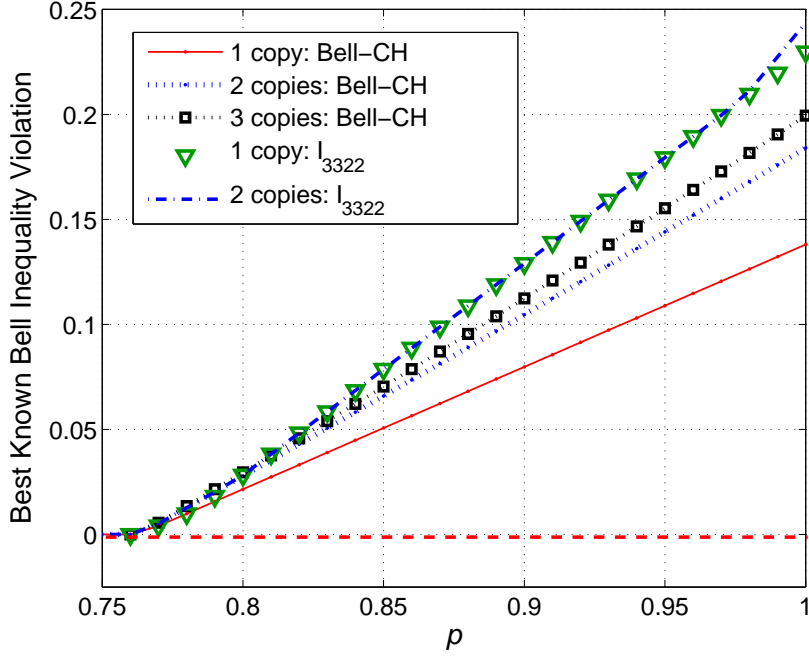


FIGURE 6.4: Best known expectation value of the Bell operator coming from the Bell-CH inequality [\mathcal{B}_{CH} , Eq. (5.39)] and the I_{3322} inequality [$\mathcal{B}_{I_{3322}}$, Eq. (5.43)], with respect to the 3-dimensional isotropic states, $\rho_{I_3}(p)$; p is the weight of maximally entangled two-qutrit state in the mixture. The single copy Bell-CH inequality violation found here through LB is identical with the maximal violation, $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho_{I_3})$, found by Ito *et al.* in Ref. [14].

6.4 Conclusion

In this chapter, we have focused on bipartite entangled systems and analyzed the extent to which a given entangled state can violate a given Bell inequality. For the Bell-CH inequality, the measurement scheme that we have presented in Sec. 6.2.1 has allowed us to obtain the best known violation of any pure two-qudit states for this inequality. A general proof that the measurement is indeed optimal seems formidable. However, the resulting violation does reproduce known (optimal) results in various special cases, including the maximal Bell-CH violation for 3-dimensional isotropic states $\rho_{I_3}(p)$. In addition, intensive numerical studies have not provided a single instance where the presented measurement is outperformed. In Sec. 6.2.2, we have also briefly reviewed the best known I_{22dd} -violation for the d -dimensional isotropic states, $\rho_{I_d}(p)$.

Next, we considered the enhancement of nonclassical correlations by collective measurements without postselection. This amounts to allowing an experiment in which a local unitary is applied to a number of copies of the state ρ prior to the Bell inequality experiment. We find that the Bell-CH inequality violation of all bipartite pure entangled states, can be enhanced by allowing collective measurements even without postselection. For mixed entangled states, however, explicit examples (Werner states) have been presented to demonstrate that there may be entangled states whose nonclassical correlations cannot be enhanced in

any Bell-CH experiments. In fact, the set of mixed two-qubit states whose Bell-CH violation can be increased with collective measurements seems to be relatively small.

We have also done some preliminary studies on how the usefulness of collective measurements depends on the choice of Bell inequality and on the dimension of the subsystem. Our data at the moment are consistent with the hypothesis that the usefulness of collective measurements in Bell inequality experiments increases with the Hilbert space dimension and with the number of measurement outcomes allowed by Bell inequality. On the other hand as the number of measurement settings allowed by the Bell inequality increases the advantage provided by collective measurements seems to diminish. However, note that we have not really performed the systematic study required to establish such trends, if they exist, due to the great numerical effort that would be required. Given these observations, it does seem that postselection is a lot more powerful than collective measurements on their own in increasing Bell-inequality violation.

An immediate question that follows from the present work is what is the class of quantum states whereby collective measurements can increase their Bell inequality violation? One motivation for studying our problem is to understand better the set of quantum states that can lead to a Bell inequality violation and are thus inconsistent with a locally causal description. It has been known for a long time that this set is a strict subset of the entangled states if projective [29] or even generalized measurements [31] on single copies of a system are permitted. One might wonder whether collective measurements without postselection allow us to violate Bell inequalities for a larger set of states. However we do not know of examples where a state that does not violate a given Bell inequality becomes violating under collective measurements when no postselection is allowed [208]. Moreover, for mixed states, the set of states whose violations increase when collective measurements are allowed appears to be rather restricted. This is consistent with the recent work by Masanes [209] which suggests that the set of states that violates a given Bell inequality under collective measurements without postselection is a subset of all distillable states.

Finally, the analysis that we have presented in this chapter only concerns bipartite quantum systems. Given that multipartite entanglement is fundamentally richer than the bipartite analogue, it should also be interesting to investigate the possibility of enhancing nonclassical correlations by collective measurements in the multipartite setting.

7

Nonstandard Bell Experiments and Hidden Nonlocality

As we have seen in Chapter 4, some quantum states, despite being entangled, *cannot* violate any Bell inequalities via a standard Bell experiment. Nonetheless, it is now well-known that nonclassical correlations can be derived from many of these entangled states if we consider more sophisticated Bell experiments which also allow appropriate local preprocessing — deriving nonclassical correlations from all entangled states via such nonstandard Bell experiments will be the subject of discussion in this chapter.

7.1 Introduction

Clearly, entanglement, being one of the most striking features offered by quantum mechanics, is in some way responsible for the generation of nonclassical correlations and hence the bizarre phenomenon of Bell inequality violation. Operationally, entanglement is defined in terms of the physical resources needed for the *preparation* of the state (c.f. Sec. 4.3.1): a multipartite state is said to be entangled if it cannot be prepared from classical correlations using local quantum operations assisted by classical communication (LOCC) [29]. This definition, however, does not tell us anything about the “behavior” of such a state. For example, is an entangled state useful in some quantum information processing task such as teleportation,¹ or does the state violate a Bell inequality? We have learned in Sec. 4.3.2 that with a standard Bell experiment, not all entangled states can violate a Bell inequality. But some of these states do violate Bell inequalities if, prior to the measurement, the state is subjected to appropriate local preprocessing. This phenomenon has been termed *hidden nonlocality* [38, 39].

¹For bipartite systems, this question has been answered in Ref. [210].

Thus far, all existing protocols that demonstrate hidden nonlocality in a nonstandard Bell experiment involve some kind of *local filtering operations*. These are local measurements that if successful are followed by a standard Bell inequality experiment, but if unsuccessful result in the state being discarded. Moreover, by allowing joint measurements on several copies of the state in conjunction with local filtering operations, Peres [40] has shown that an even larger set of two-qubit entangled states could be detected through their violation of a Bell inequality. However, the question of whether all entangled states might display some kind of (hidden) nonlocality has remained open.

A possible generalization of Peres' idea would be to perform local filtering operations and collective measurements on arbitrarily many copies of a quantum state, and subject the resulting state to a standard Bell inequality test. If the resulting correlation violates a Bell inequality, the original state is said to *violate this inequality asymptotically* [209]. In Ref. [209] it was shown that a bipartite state violates the Bell-CHSH inequality asymptotically if, and only if, it is distillable. This result suggests that undistillable entangled states may admit a locally causal description even when experiments are performed on an arbitrarily large number of copies of the state.

As a result, it does seem necessary to consider even more general protocols to derive nonclassical correlations that may be hidden in an arbitrary entangled state. One natural possibility is to allow joint processing with auxiliary states (that do not themselves violate the Bell inequality) rather than just with more copies of the state in question. In this chapter, we will show that this kind of protocol is indeed useful to derive nonclassical correlations from all entangled states. This gives a conclusive answer to the long-standing question of whether or not all entangled states can lead to observable nonlocality [31, 38, 39, 53].

The structure of this chapter is as follows. In Sec. 7.2, we will start off by reviewing some of the nonstandard Bell tests where the system of interest is measured one copy at a time. This is then followed by a more general scenario whereby collective measurements on multiple copies of the quantum system are allowed in the nonstandard Bell experiment. After that, in Sec. 7.4, we will provide a protocol involving shared ancilla states to demonstrate the nonlocality associated with all bipartite entangled states. Finally, we will conclude with some possible avenues for future research.

7.2 Single Copy Nonstandard Bell Experiments

7.2.1 Nonstandard Bell Experiments on Pure Entangled States

The very first (implicit) proposal on a nonstandard Bell experiment could be traced back to the influential work by Gisin [35]. There, he considered a general, entangled pure two-qudit state $|\Phi_d\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $d \geq 2$,

$$|\Phi_d\rangle = \sum_{i=1}^d c_i |i\rangle_A |i\rangle_B, \quad (6.1)$$

but where local measurements are performed only on an entangled two-qubit subspace.² By showing that any entangled pure two-qubit state can violate the Bell-CHSH inequality, Gisin has essentially also demonstrated that any entangled pure two-qudit state can lead to a Bell-CHSH violation by first performing the following *projections* on the local subsystems

$$\mathcal{H}_A \rightarrow \Pi_A^{(2)} \mathcal{H}_A, \quad \mathcal{H}_B \rightarrow \Pi_B^{(2)} \mathcal{H}_B, \quad (7.1)$$

where

$$\Pi_A^{(2)} \equiv |i\rangle_{AA}\langle i| + |j\rangle_{AA}\langle j|, \quad \Pi_B^{(2)} \equiv |i\rangle_{BB}\langle i| + |j\rangle_{BB}\langle j|, \quad (7.2)$$

$|i\rangle_A, |j\rangle_A \in \mathcal{H}_A$ are any pair of orthogonal basis vectors defined in Eq. (6.1) and $|i\rangle_B, |j\rangle_B$ are the corresponding correlated basis states in \mathcal{H}_B .

In effect, these local projections bring the pure two-qudit state $|\Phi_d\rangle$ into a pure two-qubit state $|\Phi_2\rangle$

$$|\Phi_d\rangle \rightarrow |\Phi_2\rangle \propto \Pi_A^{(2)} \otimes \Pi_B^{(2)} |\Phi_d\rangle, \quad (7.3)$$

with some probability of success. Clearly, such local transformation does not always succeed. In the event that it fails, the resulting state is discarded³ but whenever the transformation succeeds, the resulting two-qubit state $|\Phi_2\rangle$ is further subjected to a standard Bell-CHSH experiment to unveil its nonclassical feature. Of course, as Gisin and Peres [36] subsequently demonstrated, nonclassical correlations can also be derived directly from any pure two-qudit entangled states via a standard Bell experiment (see Sec. 6.2.1).

Whether the same can be said for multipartite entangled states still remains unclear at present. When the number of parties (denoted by n) is 3, Chen *et al.* [37] have presented strong evidence that all tripartite pure entangled states violate a Bell inequality that they have derived. Beyond this, it is still not known if a general n -partite pure entangled state can violate some Bell inequality via a standard Bell experiment. Nonetheless, as Popescu and Rohrlich showed in Ref. [202], all n -partite pure entangled states do lead to a Bell inequality violation after appropriate local filtering operations. The key idea behind their proof is to realize that by suitable choice of local projection on $n - 2$ out of n subsystems, a local transformation that brings an n -partite pure entangled state to a bipartite pure entangled state is always achievable with some nonzero probability. Then, conditioned on the success of this local transformation, the desired bipartite entangled state can further be subjected to, say, the above-mentioned scheme proposed by Gisin [35], or to the measurement scheme described in Sec. 6.2.1, which will lead to a Bell-CHSH violation coming from any n -partite pure entangled state.

7.2.2 Nonstandard Bell Experiments on Mixed Entangled States

Let us now turn our attention to mixed states. Clearly, since some mixed entangled states, e.g. the Werner states $\rho_{w_d}(p)$ with $p \leq p_{L,W}^{\text{POVM}}$, admit a general LHV description, we

²From Eq. (6.1), it is evident that any pair of correlated local bases $\{|i\rangle_A, |j\rangle_A\}, \{|i\rangle_B, |j\rangle_B\}$ would define an entangled pure two-qubit subspace for $|\Phi_d\rangle$.

³A modification to this scheme, proposed by Popescu and Rohrlich [202], would bypass postselection but, instead, perform trivial local measurements $\mathbb{1}_{d-2}$ whenever the received subsystem falls outside the qubit subspace. In this case, they showed that such measurement scheme could also lead to a (non-optimal) Bell-CHSH violation for any entangled pure two-qudit state.

cannot hope to find a Bell inequality violation of such states via a standard Bell experiment. Nonetheless, as we will see in this section, nonstandard Bell experiments — in the form of standard Bell experiments preceded with appropriate local filtering operations — can also help to demonstrate the nonlocality that is apparently *hidden* in some of these entangled quantum states.

7.2.2.1 Nonlocality Hidden in Werner States

At first glance, Werner's LHM for $\rho_{w_d}(p)$ with $p \leq p_{L,W}^\Pi$, c.f. Sec. 4.3.2.1, seems to have suggested the impossibility of deriving nonclassical correlations from such mixed entangled states. However, another nonclassical feature displayed by *all* entangled, 2-dimensional Werner states — namely, all entangled $\rho_{w_2}(p)$ were found to be useful for teleportation [52] — has led Popescu to think that there may be other less straightforward way to derive nonclassical correlations from these quantum states.

Indeed, via a nonstandard Bell experiment of the kind described in Sec. 7.2.1, Popescu [38] has managed to show that for $d \geq 5$, Werner states admitting explicit LHM can also lead to a Bell-CHSH violation. Specifically, Popescu has considered the Werner state, Eq. (4.12), with $p = p_{L,W}^\Pi$, i.e., the entangled Werner state whereby an explicit LHM for projective measurements is known. This mixture can be written more explicitly as

$$\rho_{w_d}(p_{L,W}^\Pi) = \rho_{w_d} \left(1 - \frac{1}{d} \right) = \frac{1}{d^2} \left(2\Pi_- + \frac{1}{d} \mathbf{1}_d \otimes \mathbf{1}_d \right). \quad (7.4)$$

By locally projecting each subsystems onto a 2-dimensional subspace via Eq. (7.2), i.e.,

$$\rho_{w_d}(p_{L,W}^\Pi) \rightarrow \Pi_{\mathcal{A}}^{(2)} \otimes \Pi_{\mathcal{B}}^{(2)} \rho_{w_d}(p_{L,W}^\Pi) \Pi_{\mathcal{A}}^{(2)} \otimes \Pi_{\mathcal{B}}^{(2)}, \quad (7.5)$$

and after renormalization,⁴ one obtains a 2-dimensional Werner state with $p = p' \equiv d/(d+2)$, i.e.,

$$\rho_{w_2}(p') = \frac{d}{d+2} \left(|\Psi^-\rangle\langle\Psi^-| + \frac{1}{2d} \mathbf{1}_2 \otimes \mathbf{1}_2 \right). \quad (7.6)$$

Now, if this 2-dimensional state is further subjected to local measurements that give maximal Bell-CHSH violation⁵ for the singlet state $|\Psi^-\rangle$, one finds that⁶

$$\mathcal{S}_{\text{QM}}(\rho_{w_2}(p')) = \frac{d}{d+2} \times 2\sqrt{2}, \quad (7.7)$$

which is greater than 2 for all $d \geq 5$. Therefore, for all Werner states $\rho_{w_d}(p)$ with $p = p_{L,W}^\Pi$ and $d \geq 5$, even though there exists an explicit LHM which reproduces their quantum mechanical probabilities (for projective measurements), nonclassical correlations can be still derived from them by first projecting the states locally, each onto an appropriate 2-dimensional subspace. This is an illustration of what is now commonly called *hidden nonlocality*, where the nonclassical correlations hidden in an entangled state only shows up in a more sophisticated, nonstandard Bell experiment.

⁴Note that this local transformation only succeeds with probability $\frac{2d+4}{d^3}$.

⁵This can be obtained by applying appropriate local unitary transformation to the measurement described in Eq. (6.2) and Eq. (B.21).

⁶It can be easily shown that these local measurements give zero expectation value for the maximally mixed state $\frac{\mathbf{1}_2 \otimes \mathbf{1}_2}{4}$.

7.2.2.2 Nonlocality Hidden in Standard Bell-CHSH Experiment

As opposed to Popescu's example, the term, "hidden nonlocality" has also been used in a looser sense where non-Bell-CHSH-inequality-violating quantum states become Bell-CHSH-inequality-violating in a nonstandard Bell experiment [39]. In Ref. [39], Gisin has considered a class of two-qubit states that is local unitarily equivalent to

$$\rho_G(p, \theta) \equiv p |\Phi_2\rangle\langle\Phi_2| + \frac{1}{2}(1-p) \left(|0\rangle_{\mathcal{A}\mathcal{A}}\langle 0| \otimes |1\rangle_{\mathcal{B}\mathcal{B}}\langle 1| + |1\rangle_{\mathcal{A}\mathcal{A}}\langle 1| \otimes |0\rangle_{\mathcal{B}\mathcal{B}}\langle 0| \right), \quad (7.8)$$

where $p_0 \equiv 1/(2 - \sin 2\theta) < p < 1$. This state can be interpreted as a mixture of the non-maximally entangled pure-two qubit state $|\Phi_2\rangle$, c.f. Eq. (6.14), and the pure product states $|0\rangle_{\mathcal{A}}|1\rangle_{\mathcal{B}}$, $|1\rangle_{\mathcal{A}}|0\rangle_{\mathcal{B}}$. Using the PPT criterion for separability [161, 162], it can be easily shown that this mixture represents an entangled state whenever $p > p_E \equiv 1/(1 + \sin 2\theta)$.

Moreover, from the Horodecki criterion (Appendix B.3.2), it is also not difficult to show that despite being entangled, $\rho_G(p, \theta)$ with

$$p_E < p \leq p_L \equiv \frac{4}{4 + \sin^2 2\theta}, \quad (7.9)$$

does not violate the standard Bell-CHSH inequality with *any* choice of dichotomic measurements. However, in practice, even if the source emits physical systems that are well described by $\rho_G(p, \theta)$, it is not inconceivable that the *end users* Alice and Bob may receive states that are better described by a different density matrix $\rho'_G(p, \theta)$, which could well lead to a Bell inequality violation.

In particular, if $\rho_G(p, \theta)$ describes the polarization state of photon pairs emitted from some source and where each pair of photons are distributed, respectively, to Alice and Bob via channels that both perform the following local filtering operation

$$F_{\mathcal{A}} = F_{\mathcal{B}} = \begin{pmatrix} \sqrt{\tan \theta} & 0 \\ 0 & 1 \end{pmatrix}, \quad (7.10)$$

then, at the end of the channels, Alice and Bob will receive a state that is actually better described by

$$\rho'_G(p, \theta) \propto F_{\mathcal{A}} \otimes F_{\mathcal{B}} \rho_G(p, \theta) F_{\mathcal{A}}^\dagger \otimes F_{\mathcal{B}}^\dagger.$$

More explicitly, after normalization, the locally filtered state reads

$$\rho'_G(p, \theta) = \frac{\tan \theta}{p_{\text{suc.}}} \left[p \sin 2\theta |\Phi_2^+\rangle\langle\Phi_2^+| + \frac{1}{2}(1-p) \left(|0\rangle_{\mathcal{A}\mathcal{A}}\langle 0| \otimes |1\rangle_{\mathcal{B}\mathcal{B}}\langle 1| + |1\rangle_{\mathcal{A}\mathcal{A}}\langle 1| \otimes |0\rangle_{\mathcal{B}\mathcal{B}}\langle 0| \right) \right],$$

where $p_{\text{suc.}} = \tan \theta [1 - p(1 - \sin 2\theta)]$ is the probability that they both receive a photon at their end. Note that in contrast with the original state given by Eq. (7.8), the resulting state $\rho'_G(p, \theta)$ can now be described as a mixture of the maximally entangled pure two-qubit state $|\Phi_2^+\rangle$ and the same set of pure product states involved in Eq. (7.8).

Again, from the Horodecki criterion, it can be shown that the locally filtered state $\rho'_G(p, \theta)$ violates the Bell-CHSH inequality if and only if

$$p > p'_L \equiv \frac{1}{1 + (\sqrt{2} - 1) \sin 2\theta}. \quad (7.11)$$

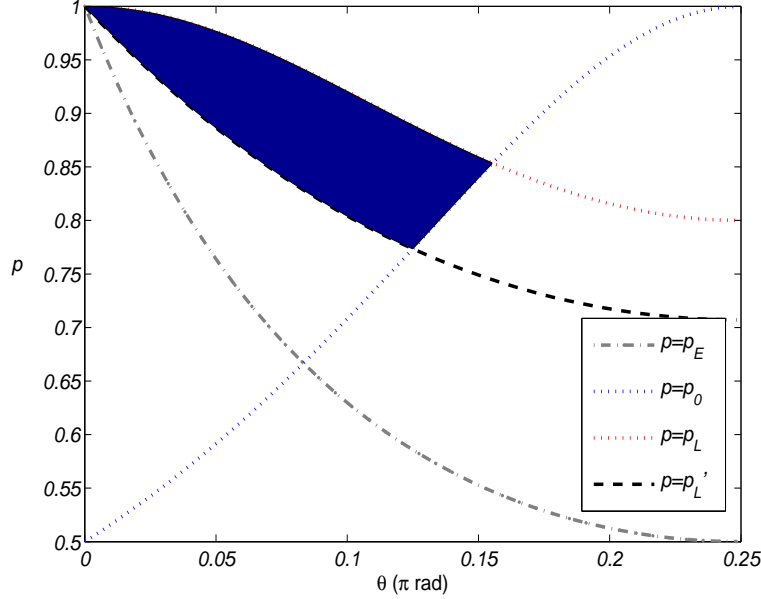


FIGURE 7.1: The relevant parameter space for $\rho_G(p, \theta)$. The set of states that do not violate the Bell-CHSH inequality but which do after the local filtering operations given by Eq. (7.10) is the shaded region bounded by the black dashed line ($p = p'_L$), the blue dotted line ($p = p_0$) and the red dotted line ($p = p_L$).

Now, if the intersection of the sets satisfying $p > p_0$, Eq. (7.9) and Eq. (7.11) is not empty, one will have found example(s) of two-qubit state not violating the Bell-CHSH inequality but which does after the local filtering operations given by Eq. (7.10). Indeed, as can be seen from Figure 7.1, a substantial subset of the class of states $\rho_G(p, \theta)$ do satisfy the conjunction of all the above requirements. Hence, as was first shown by Gisin [39], there are two-qubit states whose nonclassical correlations cannot be observed directly in a standard Bell-CHSH experiment but if the experiment is preceded with appropriate local filtering operations, their hidden nonlocality do lead to observable nonclassical correlations. It is worth noting that an experimental demonstration of a very similar example has been carried out and presented in Ref. [28].

7.2.3 Justification of Single Copy Nonstandard Bell Experiment

As we have seen in the examples given above, even if a bipartite entangled quantum state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is not known to violate a Bell inequality (or in some cases is known to be NBIV), it may still be possible to observe a Bell inequality violation coming from ρ , if, prior to the standard Bell experiment, appropriate local filtering operations are carried out. In effect, this transforms the state ρ locally to another quantum state $\Omega(\rho)$ via:

$$\rho \rightarrow \Omega(\rho) = \sum_i F_{A,i} \otimes F_{B,i} \rho F_{A,i}^\dagger \otimes F_{B,i}^\dagger, \quad (7.12)$$

where $F_{\mathcal{A},i}$ and $F_{\mathcal{B},i}$ are, respectively, local filtering operators (*aka* Kraus operators⁷) acting on subsystem \mathcal{A} and \mathcal{B} . Up to some constant, $\Omega(\rho)$ is also known as a *separable map* acting on ρ . Evidently, since these local transformations do not always succeed, some form of postselection, and hence (classical) communication is involved when transforming the state locally from ρ to $\Omega(\rho)$. Indeed, this nondeterministic nature of the local operations also result in them being more commonly known in the literature as *stochastic local quantum operations assisted by classical communication* (henceforth abbreviated as SLOCC, Appendix A.2) [213].

Naturally, the postselection involved in such SLOCC prior to a standard Bell experiment reminds one of the *detection loophole* discussed in a standard Bell test. An important distinction between the two, as was first pointed out by Popescu [38], and subsequently by Żukowski *et al.* [147], is that the postselection is carried out *prior to*⁸ the standard Bell experiment. Therefore, *a priori*, the postselection involved does not causally depend on the choice of measurements made subsequently. In addition, one should note that local filtering operation on any quantum state ρ *cannot* create nonclassical correlations in the resulting state ρ' — local operations assisted by classical communication cannot create entanglement. As such, any nonclassical correlations derivable from the resulting state ρ' must have inherited from the original state ρ . For a more rigorous version of this argument, see the proof presented by Żukowski and coauthors in Ref. [147].

7.3 Nonstandard Bell Experiments on Multiple Copies

The single-copy nonstandard Bell experiments that we have considered in the previous section has certainly shed some light on what can be done to reveal the nonclassical correlations associated with an entangled quantum system. A natural question that follows is whether this aspect of nonclassicality can be demonstrated for arbitrary entangled states. To this end, a negative answer was provided by Verstraete and Wolf [195] who showed that a large class of two-qubit entangled states, including some of the entangled Werner states, do not violate the Bell-CHSH inequality even after an arbitrary local filtering operation.

Of course, as with the complication involved in a standard Bell experiment, it is still possible, at least in principle, that some of these states actually violate some other more complicated Bell inequalities after appropriate SLOCC. Nevertheless, given that not much is known in this regard — even in the simpler scenario of a standard Bell experiment — it seems natural to consider other alternatives. In particular, one could consider running a standard Bell experiment using collective measurements on multiple copies of a quantum system. The idea is that perhaps, one can find a quantum state ρ not known to violate *any* Bell inequality when measured one copy at a time but for N large enough, one finds that $\rho^{\otimes N}$ does violates some Bell inequality. However, as we have discussed in Chapter 6 (see Sec. 6.3.2 in particular) no such example has ever been found.

⁷After Kraus' work on completely positive maps [211, 212]. Eq. (7.12) is also commonly known in the literature as the Kraus decomposition of $\Omega(\rho)$.

⁸In the relativistic sense.

7.3.1 Nonstandard Bell Experiments with Collective Measurements

In the same vein as the single-copy scenario, why not consider a standard Bell experiment that is preceded with SLOCC on multiple copies of a quantum system? More precisely, even when ρ , as well as $\rho^{\otimes N}$ is not found to violate any Bell inequality, it could still be that the following local filtering operations prior to a standard Bell experiment is useful in deriving nonclassical correlations from $\rho^{\otimes N}$:

$$\rho^{\otimes N} \rightarrow \rho' \propto \sum_i F_{A,i} \otimes F_{B,i} \rho^{\otimes N} F_{A,i}^\dagger \otimes F_{B,i}^\dagger, \quad (7.13)$$

where here, it is worth noting that the tensor product between $F_{A,i}$ and $F_{B,i}$ acts differently from the tensor product involved in $\rho^{\otimes N}$.

Indeed, this is exactly what Peres has contemplated to demonstrate the nonlocality hidden in 2-dimensional Werner states [40]. More specifically, Peres has considered a scenario where N copies of $\rho_{W_2}(p)$ are collected and further subjected to some local unitary transformation acting on all the N copies of the local subsystems. After that, for both Alice and Bob, projective measurements are carried out in the Z basis for all but one of the N particles.⁹ If all the $2(N-1)$ measurement results are “ \uparrow ”, the remaining 2 particles are subjected to a standard Bell-CHSH experiment, otherwise they are discarded and the experiment is restarted.

With this protocol, Peres has shown that many $\rho_{W_2}(p)$ not known to violate any Bell inequality do violate the Bell-CHSH inequality after the described postselection. In particular, with $N = 5$ copies, Peres has found that, despite the explicit LHV constructed by Werner (see Sec. 4.3.2.1), the Werner state $\rho_{W_2}(1/2)$ does lead to a Bell-CHSH inequality violation of 2.0087 via the above-mentioned nonstandard Bell experiment. Moreover, due to the distillability [164] of all 2-dimensional entangled states [214], it is expected that for sufficiently large N , all entangled $\rho_{W_2}(p)$ will lead to a Bell-CHSH inequality violation in this manner.

7.3.2 Nonstandard Bell Experiments and Distillability

As we have just seen, a nonstandard Bell experiment that allows collective measurement on many copies of a quantum system and postselection on some desired outcome is clearly more powerful than all the other Bell experiments that we have described so far. In particular, if we allow N — the number of copies — to be arbitrarily large, it seems like we can go through these procedures to derive nonclassical correlations out of a large set of entangled states. The immediate question that follows is whether this is a *strict subset* of the set of entangled states. Evidently, if a state ρ is distillable, one can extract a spin-1/2 singlet state from $\rho^{\otimes N}$ via some local filtering operations, c.f. Eq. (7.13), and therefore ρ violates a standard Bell-CHSH experiment that is preceded with some SLOCC.

What about the converse? Must *undistillable entangled states* (aka *bound entangled states*) satisfy Bell inequalities even if the Bell experiment is preceded with arbitrary SLOCC? To answer this question, Masanes has introduced the following definition in Ref. [209].

⁹If there is a need to perform measurement in any other basis, one can achieve that by first performing additional unitary transformation on the particle in question prior to a measurement on the Z basis [40].

Definition 12. A quantum state ρ is said to violate a Bell inequality asymptotically if $\rho^{\otimes N}$ for an arbitrarily large N violates the Bell inequality after some stochastic local quantum operations without communication (SLO).

Notice that no communication is allowed in the above definition. However, as it turns out, allowing classical communication (i.e., with SLOCC instead of SLO) does not allow more states to violate a Bell inequality in this manner [209].¹⁰ A partial result to the above question is then provided by Masanes in the following theorem [209].

Theorem 13. A bipartite state ρ is distillable if and only if it violates the Bell-CHSH inequality asymptotically. In other words ρ is distillable if and only if there exists an $N \in \mathbb{Z}^+$ and some SLO, denoted by Ω such that $\Omega(\rho^{\otimes N})$ violates the Bell-CHSH inequality.

Again, it is still logically possible that undistillable states can violate some other Bell inequalities asymptotically. Nonetheless, this theorem due to Masanes has clearly suggested that one should also look for other alternatives to derive nonclassical correlations, if any, associated with arbitrary bipartite entangled states, especially the bound entangled states.

7.4 Observable Nonlocality for All Bipartite Entangled States

In this section, we will go beyond the typical nonstandard Bell experiment and consider one that also involves shared ancilla states. In particular, we will prove that via a local filtering protocol that involves a specific ancilla state (which by itself does not violate the Bell-CHSH inequality), one can always observe a Bell-CHSH violation coming from a single copy of any bipartite entangled state.

7.4.1 Bipartite States with No Bell-CHSH Violation after SLOCC

To this end, let us first introduce the following definition regarding the set of bipartite states that *do not violate* the Bell-CHSH inequality even after arbitrary local filtering operations. The nonstandard Bell experiment that we are going to consider will involve an ancilla state which is a member of this set.

Definition 14. Denote by $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ the set of bipartite states that do not violate the Bell-CHSH inequality, even after SLO on a single copy of the state of interest.

As for Theorem 13, it follows from the results presented in Ref. [209] that states in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ also do not violate the Bell-CHSH inequality even after SLOCC — hence the notation $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. The exact nature of the local operations allowed in the definition of $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ is thus not important. Clearly, states that do not violate the Bell-CHSH inequality asymptotically, c.f. Definition 12, are in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Therefore, it follows from Theorem 13 that $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ contains all undistillable states [209] (which include the set of bound entangled states as a subset).

¹⁰Intuitively, one can see that this is true by noting that the role of classical communication, if any, in a nonstandard Bell experiment is primarily to facilitate any postselection involved.

As remarked earlier, there are no undistillable two-qubit entangled states [214]. However, from the results presented in Ref. [195], we know that there are also two-qubit entangled states that are in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$.

In what follows, we will describe a set of necessary and sufficient conditions for a general bipartite state ρ to be in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. To begin with, we note that $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ is a *convex set*¹¹ and thus it can be characterized via hyperplanes that separate this set from any point outside the set. In particular, for any state ρ that is not in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, a hyperplane that separates ρ from $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ can be constructed; this hyperplane therefore serves as a kind of witness operator that detects Bell-CHSH violation of ρ after some SLOCC.

Lemma 15. *A bipartite state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ belongs to $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ if, and only if, it satisfies*

$$\text{tr} [\rho (F_A \otimes F_B)^\dagger H_\theta (F_A \otimes F_B)] \geq 0, \quad (7.14)$$

for all matrices of the form $F_A: \mathcal{H}_A \rightarrow \mathbb{C}^2$, $F_B: \mathcal{H}_B \rightarrow \mathbb{C}^2$ and all $\theta \in [0, \pi/4]$, where

$$H_\theta \equiv \mathbb{1}_2 \otimes \mathbb{1}_2 - \cos \theta \sigma_x \otimes \sigma_x - \sin \theta \sigma_z \otimes \sigma_z, \quad (7.15)$$

$\mathbb{1}_2$ being the 2×2 identity matrix and $\{\sigma_i\}_{i=x,y,z}$ are the Pauli matrices introduced in Eq. (2.6).

Proof. We shall prove this Lemma in two stages. Firstly, we will prove a criterion analogous to inequality (7.14) for the scenario where no local filtering operation is involved and when ρ is a two-qubit state. Then, we will provide a proof for the general scenario by incorporating existing results in Ref. [209].

Now, let us start with the special case of a two-qubit state and where no local filtering operation is involved. Recall from Sec. 5.3.2 that in a standard Bell experiment — a Bell experiment without local preprocessing — a two-qubit state ϱ violates the Bell-CH/ Bell-CHSH inequality if and only if [30, 195]

$$\varsigma_1^2 + \varsigma_2^2 > 1, \quad (5.42)$$

where ς_k is the k^{th} largest singular value of the 3×3 real matrix T defined in Eq. (5.38). Equivalently, $(\varsigma_1, \varsigma_2)$ derived from ϱ must lie outside the unit circle $\varsigma_1^2 + \varsigma_2^2 = 1$, which is true if and only if there exists $\theta \in [0, 2\pi]$ such that

$$\varsigma_1 \cos \theta + \varsigma_2 \sin \theta > 1. \quad (7.16)$$

Now, it is also well-known that by appropriate local unitary transformations U, V , it is always possible to arrive at a local basis such that T is diagonal¹² with $\varsigma_1 = T_{xx}$ and $\varsigma_2 = T_{zz}$. From the definition of T it then follows that

$$\varsigma_1 \cos \theta = \text{tr} [(U \otimes V) \varrho (U \otimes V)^\dagger (\cos \theta \sigma_x \otimes \sigma_x)], \quad (7.17)$$

with the expression for $\varsigma_2 \sin \theta$ involving obvious modifications. Since singular values are non-negative, it thus follows that if ϱ violates the Bell-CHSH inequality then there exist $U, V \in \text{SU}(2)$, $\theta \in [0, \frac{\pi}{4}]$ such that

$$\text{tr} [\varrho (U \otimes V)^\dagger H_\theta (U \otimes V)] < 0. \quad (7.18)$$

¹¹The proof is similar to the one presented in Appendix B.2.1.

¹²See for example pp. 2227 of Ref. [215].

Conversely, suppose that there exists some $U, V \in \text{SU}(2)$, $[0, \frac{\pi}{4}]$ satisfying inequality (7.18), then it follows that

$$T_{xx} \cos \theta + T_{zz} \sin \theta > 1. \quad (7.19)$$

Since $\varsigma_1 \geq \varsigma_2$ by definition, the inequalities

$$|T_{ii}| \leq \varsigma_1 \leq 1, \quad i \in \{x, y, z\}, \quad (7.20)$$

follow from the definition of singular values [216] and the well-known fact that all singular values of T are less than or equal to one.¹³ In addition, since $0 \leq \theta \leq \frac{\pi}{4}$, we must also have

$$\cos \theta \geq \sin \theta \geq 0.$$

These inequalities, together with Eq. (7.19) and Eq. (7.20), imply that both T_{xx} and T_{zz} must be non-negative. Moreover, we may assume without loss of generality that $T_{xx} \geq T_{zz}$. This is because if it happens that $T_{xx} = \min\{T_{xx}, T_{zz}\}$, then since

$$T_{zz} \cos \theta + T_{xx} \sin \theta \geq T_{xx} \cos \theta + T_{zz} \sin \theta > 1,$$

we may also take the larger of $\{T_{xx}, T_{zz}\}$ as the coefficient of $\cos \theta$. Finally, note that the singular values of T obey the inequality $|T_{xx} + T_{zz}| \leq \varsigma_1 + \varsigma_2$ (pp. 76, Ref. [216]). As a result, we find

$$\begin{aligned} \varsigma_1 \cos \theta + \varsigma_2 \sin \theta &= \varsigma_1 (\cos \theta - \sin \theta) + (\varsigma_1 + \varsigma_2) \sin \theta, \\ &\geq \varsigma_1 (\cos \theta - \sin \theta) + (T_{xx} + T_{zz}) \sin \theta, \\ &\geq T_{xx} \cos \theta + T_{zz} \sin \theta > 1, \end{aligned}$$

so ρ violates the Bell-CHSH inequality. Thus, a two-qubit state ρ *violates the Bell-CHSH inequality if and only if inequality (7.18) holds*. This completes our proof for the scenario where no local filtering operation is involved and when ρ is a two-qubit state.

Let us now come back to the question of Bell-CHSH violation after local filtering operations. Assume that ρ violates Bell-CHSH inequality after SLO. Let us show that it must violate inequality (7.14) for some (F_A, F_B, θ) . In Ref. [209], it was proven that, if a state violates the Bell-CHSH inequality, then it can be transformed by SLO into a two-qubit state which also violates the Bell-CHSH inequality. Therefore, there must exist a separable map Ω with two-qubit output, such that the resulting state $\varrho = \Omega(\rho)$ satisfies inequality (7.18) for some (U, V, θ) which we shall denote by (U_0, V_0, θ_0) , i.e.,

$$\text{tr} [\Omega(\rho) (U_0 \otimes V_0)^\dagger H_{\theta_0} (U_0 \otimes V_0)] < 0,$$

Clearly, if this is true, it also follows from the Kraus decomposition of $\Omega(\rho)$, Eq. (7.12), such that

$$\text{tr} \left[\left(F_{A,i} \otimes F_{B,i} \rho F_{A,i}^\dagger \otimes F_{B,i}^\dagger \right) (U_0 \otimes V_0)^\dagger H_{\theta_0} (U_0 \otimes V_0) \right] < 0,$$

¹³See, for example, pp. 1840 of Ref. [217].

for some i . This implies that ρ violates inequality (7.14) for $F_{\mathcal{A}} = U_0 F_{\mathcal{A},i}$, $F_{\mathcal{B}} = V_0 F_{\mathcal{B},i}$ and $\theta = \theta_0$. This proves one direction of the lemma, we shall next show the other.

Assume that ρ violates inequality (7.14) for $(F_{\mathcal{A},0}, F_{\mathcal{B},0}, \theta_0)$. It is straightforward to see that ρ violates the Bell-CHSH inequality after SLOCC. Consider the operation that transforms ρ into the two-qubit state $\varrho \propto (F_{\mathcal{A},0} \otimes F_{\mathcal{B},0}) \rho (F_{\mathcal{A},0} \otimes F_{\mathcal{B},0})^\dagger$. By assumption, the final state ϱ satisfies inequality (7.18) with $U = V = \mathbb{1}_2$ and $\theta = \theta_0$, which implies that it violates the Bell-CHSH inequality. This completes our proof of the Lemma. \square

7.4.2 Nonstandard Bell Experiment with Shared Ancillary State

With the characterization given above, we are now ready to state and prove the main result of this section, namely:

Theorem 16. *A bipartite state σ is entangled if, and only if, there exists a state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that $\rho \otimes \sigma$ is not in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$.*

Let us first try to clarify the physical significance behind this theorem. If ρ belongs to $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, no matter how much additional classical correlation (which can always be represented by a separable state η_{sep}) we supply to it, the result $\rho \otimes \eta_{\text{sep}}$ is still in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. On the contrary, for every entangled state σ , we can always find a $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that the combined state $\rho \otimes \sigma$ is not in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, and hence violates the Bell-CHSH inequality after appropriate SLOCC. *This is true, remarkably, even if both ρ and σ are in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$.*

Here, the violation of Bell-CHSH inequality manifests the qualitatively different behavior between $\rho \otimes \sigma$ and $\rho \otimes \eta_{\text{sep}}$, where η_{sep} is any separable state, and σ is any entangled state. In other words, Theorem 16 says that for each entangled state σ there exists a protocol (which also involves the ancilla state ρ associated with the theorem) in which σ cannot be substituted by an arbitrarily large amount of classical correlations without changing the experimental statistics.¹⁴ Consequently, yet another way of putting the theorem would be: *bipartite entangled states are the ones that cannot always be simulated by classical correlations.*

The proof of the above theorem relies on an explicit characterization of the set $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, which we have already obtained in Sec. 7.4.1. We can then make use of convexity arguments similar to those given in Ref. [210] to prove by contradiction that there exists some $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that one of those witness-like operators may be constructed for $\rho \otimes \sigma$ whenever σ is entangled. To carry this argument through we also require a characterization of the separable completely positive maps between Bell diagonal states, which we have included in Appendix A. With these characterizations in hand, we may now proceed to the actual proof of the theorem.

Proof. Firstly, we note that if σ is separable, then for all $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ we must have $\rho \otimes \sigma \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Intuitively, one can see that this is so because σ can only generate classical correlations which will not lead to any Bell inequality violation. In fact, starting from ρ , one can prepare $\rho \otimes \sigma$ for any separable σ during the LOCC preprocessing of ρ . Therefore,

¹⁴On the contrary, recall from the discussion in Sec. 3.1 that the existence of an LHVM for some experimental data ensures that the latter can be replaced by classical correlations which do preserve the experimental statistics given by all the joint and marginal probabilities.

if $\rho \otimes \sigma$ for any separable σ were to violate Bell-CHSH inequality after SLOCC, so would ρ , which contradicts our assumption that $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Hence, *if there exists a bipartite state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that $\rho \otimes \sigma \notin \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, we know that σ has to be entangled.* Next, we will proceed to prove the other direction of the theorem, namely

$$\sigma \text{ is entangled} \Rightarrow \exists \rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})} \text{ such that } \rho \otimes \sigma \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$$

Denote by $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ the state space that σ acts on and by d_A, d_B , respectively, the dimension of the local subsystem \mathcal{H}_A and \mathcal{H}_B . From now onwards, we will assume that σ is entangled across \mathcal{H}_A and \mathcal{H}_B . Our goal is to show for every σ , there always exists an ancilla state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that $\rho \otimes \sigma \notin \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. To achieve that, we will consider ancilla state ρ that acts on the bipartite Hilbert space $[\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}] \otimes [\mathcal{H}_{B'} \otimes \mathcal{H}_{B''}]$, where $\mathcal{H}_{A'} = \mathcal{H}_A$, $\mathcal{H}_{B'} = \mathcal{H}_B$ and $\mathcal{H}_{A''} = \mathcal{H}_{B''} = \mathbb{C}^2$ (see Figure 7.2). To prove the above theorem, we then need to show that the state $\rho \otimes \sigma$ violates inequality (7.14) for some choice of F_A, F_B , and θ .

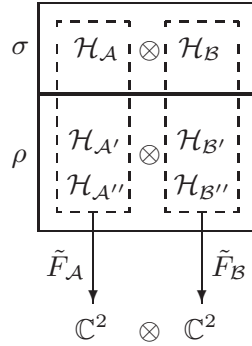


FIGURE 7.2: Schematic diagram illustrating the local filtering operations \tilde{F}_A and \tilde{F}_B involved in our protocol. The solid box on top is a schematic representation of the state σ whereas that on the bottom is for the ancilla state ρ . Left and right dashed boxes, respectively, enclose the subsystems possessed by the two experimenters \mathcal{A} and \mathcal{B} .

In particular, let

$$\tilde{F}_A = \langle \Phi_{AA'} | \otimes \mathbb{1}_{A''}, \quad \tilde{F}_B = \langle \Phi_{BB'} | \otimes \mathbb{1}_{B''}, \quad \theta = \frac{\pi}{4}, \quad (7.21)$$

where $|\Phi_{AA'}\rangle = \sqrt{d_A} |\Phi_{d_A}^+\rangle$, c.f. Eq. (4.17), is the (unnormalized) maximally entangled state between the spaces \mathcal{H}_A and $\mathcal{H}_{A'}$ (which have the same dimension), and $\mathbb{1}_{A''}$ is the identity matrix acting on \mathbb{C}^2 (analogously for Bob). After some simple calculations, it can be shown that for any ρ acting on $[\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}] \otimes [\mathcal{H}_{B'} \otimes \mathcal{H}_{B''}]$

$$\text{tr} \left[\rho \otimes \sigma (\tilde{F}_A \otimes \tilde{F}_B)^\dagger H_{\frac{\pi}{4}} (\tilde{F}_A \otimes \tilde{F}_B) \right] = \text{tr} \left[\rho (\sigma^T \otimes H_{\frac{\pi}{4}}) \right],$$

where σ^T stands for the transpose of σ .¹⁵ Hence, the requirement that inequality (7.14) is violated (i.e., $\rho \otimes \sigma \notin \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$) with $\theta = \pi/4$, $F_A = \tilde{F}_A$, $F_B = \tilde{F}_B$ becomes

$$\text{tr} \left[\rho (\sigma^T \otimes H_{\frac{\pi}{4}}) \right] < 0. \quad (7.22)$$

¹⁵More generally, for any ρ acting on $[\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}] \otimes [\mathcal{H}_{B'} \otimes \mathcal{H}_{B''}]$, any σ acting on $[\mathcal{H}_A] \otimes [\mathcal{H}_B]$, and any

What remains is to show that there exists some physical state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that the above inequality holds true.

For convenience, in the rest of the proof we allow ρ to be unnormalized. The only constraints on the matrices $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ are then positive semidefiniteness ($\rho \in \mathcal{S}^+$), and satisfiability of all the inequalities (7.14) in Lemma 15. $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ is now a convex cone, and its dual cone is defined as

$$\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH}) *} = \{X : \text{tr}(\rho X) \geq 0, \forall \rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}\}, \quad (7.23)$$

where X are Hermitian matrices. An important point to note now is that Farkas' Lemma [218] states that all matrices in $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH}) *}$ can be written as non-negative linear combinations of matrices $P \in \mathcal{S}^+$ and matrices of the form

$$(F_{\mathcal{A}} \otimes F_{\mathcal{B}})^{\dagger} H_{\theta} (F_{\mathcal{A}} \otimes F_{\mathcal{B}}) \quad (7.24)$$

with $F_{\mathcal{A}} : \mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{A}''} \rightarrow \mathbb{C}^2$ and $F_{\mathcal{B}} : \mathcal{H}_{\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{B}''} \rightarrow \mathbb{C}^2$.

We now show that there always exists $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ satisfying inequality (7.22) by supposing otherwise and arriving at a contradiction. Suppose that *for all* $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$, the converse inequality of Eq. (7.22) holds true, i.e.,

$$\text{tr}[\rho(\sigma^T \otimes H_{\frac{\pi}{4}})] \geq 0. \quad (7.25)$$

It then follows from the definition of $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH}) *}$, Eq. (7.23), that the matrix $\sigma^T \otimes H_{\frac{\pi}{4}}$ belongs to $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH}) *}$. Applying Farkas' Lemma [218] we can write

$$\sigma^T \otimes H_{\frac{\pi}{4}} = \int dx (F_{\mathcal{A},x} \otimes F_{\mathcal{B},x})^{\dagger} H_{\theta_x} (F_{\mathcal{A},x} \otimes F_{\mathcal{B},x}) + \int dy P_y,$$

where x is a label for matrices of the form given by Eq. (7.24) and y is a label for element in \mathcal{S}^+ . It is easy to see that the above matrix equality is equivalent to the matrix inequality

$$\sigma^T \otimes H_{\frac{\pi}{4}} - \int dx \Omega_x(H_{\theta_x}) \geq 0, \quad (7.26)$$

where each Ω_x is a separable map, c.f. Eq. (7.12), that takes matrices acting on $[\mathbb{C}^2] \otimes [\mathbb{C}^2]$ to matrices acting on $[\mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{A}''}] \otimes [\mathcal{H}_{\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{B}''}]$. The following Lemma, however, requires that this is true only if σ is separable (see Appendix B.5.1 for details).

Lemma 17. *Let $\Omega_x : [\mathbb{C}^2] \otimes [\mathbb{C}^2] \rightarrow [\mathcal{H}_{\mathcal{A}} \otimes \mathbb{C}^2] \otimes [\mathcal{H}_{\mathcal{B}} \otimes \mathbb{C}^2]$ be a family of maps, separable with respect to the partition denoted by the brackets. Let μ be a unit-trace, PSD matrix acting on $[\mathcal{H}_{\mathcal{A}}] \otimes [\mathcal{H}_{\mathcal{B}}]$ such that*

$$\mu^T \otimes H_{\frac{\pi}{4}} - \int dx \Omega_x(H_{\theta_x}) \geq 0, \quad (7.27)$$

where H_{θ} is defined in Eq. (7.15), then μ has to be separable.

M acting on $[\mathcal{H}_{\mathcal{A}''}] \otimes [\mathcal{H}_{\mathcal{B}''}]$, it can be shown that

$$\text{tr}[\rho \otimes \sigma (\tilde{F}_{\mathcal{A}} \otimes \tilde{F}_{\mathcal{B}})^{\dagger} M (\tilde{F}_{\mathcal{A}} \otimes \tilde{F}_{\mathcal{B}})] = \text{tr}[\rho(\sigma^T \otimes M)].$$

Hence, if all $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ are such that none of them can give rise to a Bell-CHSH violation for $\rho \otimes \sigma$ via the protocol given in Eq. (7.21) (see also Figure 7.2), it must be the case that σ is a separable state. As a result, the corresponding contrapositive positive statement reads: *for every entangled σ , there exists $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that $\rho \otimes \sigma$ violates the Bell-CHSH inequality via the protocol given by Eq. (7.21)*. This completes our proof of Theorem 16. \square

At this stage, it is worth making a few other remarks concerning the nonstandard Bell experiments that we have just described. To fix ideas, we will restrict ourselves to the nontrivial case that both ρ and σ are members of $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ and where σ is entangled. Then for $\rho \otimes \sigma$ to violate the Bell-CHSH inequality via our protocol, it must also be that (1) ρ is entangled and (2) at least one of ρ and σ has negative partial transposition. That ρ is entangled can be easily seen by following the argument given in the proof of Theorem 16, but with the role of ρ and σ reversed (pp. 96). On the other hand, it is also not difficult to see that if both ρ and σ were to have positive partial transposition, then after the local filtering operation given by Eq. (7.21), the resulting two-qubit state would still be PPT [163] and hence separable [162]. Since no separable state can violate the Bell-CHSH inequality, at least one of ρ and σ must have negative partial transposition.

Meanwhile, we have only required that the ancilla state ρ does not violate the Bell-CHSH inequality, and therefore it may violate other Bell inequalities, like anyone among the zoo of inequalities presented in Sec. 3.3.1 and Sec. 3.3.2. However, even if ρ does violate another Bell inequality, we know by definition that ρ , and thus $\rho \otimes \eta_{\text{sep}}$ (with η_{sep} being any separable state) does not violate the Bell-CHSH inequality. Hence, in the Bell-CHSH experiment that we are considering, σ cannot be replaced by any classical correlations or separable state η_{sep} .

7.5 Conclusion

In this chapter, we have reviewed the phenomenon of *hidden nonlocality* associated with entangled states, and the various kinds of nonstandard Bell experiments that have been proposed to derive nonclassical correlations from them. To date, it is still not known if all entangled states can violate some Bell inequalities via a nonstandard Bell experiment that only involves the state in question. Given this state of affair, we have looked into the possibility of deriving nonclassical correlations from all bipartite entangled states by considering nonstandard Bell experiments that also involve shared auxiliary states. Evidently, the choice of such an ancilla state cannot be arbitrary. In particular, the protocol that we have considered involves an ancilla state ρ which by itself *does not violate* the Bell-CHSH inequality even after arbitrary local filtering operations. In the notation that we have developed, we say that $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Then, by considering a specific local filtering protocol, we have shown that for every entangled state σ , there exists an ancilla state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that the combined state $\rho \otimes \sigma$ *does violate* the Bell-CHSH inequality after the prescribed local filtering operations.

This provides us with a new way to interpret (bipartite) entanglement in terms of the behavior of the states, in contrast with the usual definition in terms of the preparation of the states. Entangled states are, by definition, the ones that cannot be generated from classical

correlations using LOCC. We have shown that in the bipartite case, one can equivalently define entangled states as the ones that cannot be simulated by classical correlations alone.¹⁶ In addition, this also gives a conclusive answer to the long-standing question of whether all (bipartite) entangled states can display some hidden nonlocality [31, 38, 39, 53].

Despite that, it is worth reminding that our proof of the key result is a non-constructive one. Therefore, even though we know that there exists some ancilla state ρ such that $\rho \otimes \sigma$ can lead to observable nonlocality for any entangled σ , we do not know much about the property of the ancilla state. A natural task that follows from our findings is thus to obtain an explicit expression for the ancilla state ρ for some given σ . From an experimental point of view, a better understanding of this ancilla state ρ is also relevant, since distillation protocol involving many copies of the same quantum system is hard to implement. Therefore, a protocol to demonstrate nonclassical correlations involving only a single copy of ρ and σ may be preferable over those other which involve, say, 10 copies of σ or ρ .

On the other hand, as with the bipartite scenario, there are also mixed multipartite entangled states that admit explicit LHV for projective measurement [32] (see Sec. 4.3.2.3). An interesting question that follows from the present work is therefore to determine if the current proof of observable nonlocality also generalizes to this more complicated scenario, and hence establishes some kind of equivalence between entanglement and states that cannot always be simulated by classical correlations.

¹⁶However, some nonclassical correlations can be simulated by classical correlations when supplemented with only one bit of classical communication (see, for example Ref. [219] and references therein).

8

Conclusion

It is one of the most phenomenal discoveries that quantum mechanical predictions on entangled, spatially separated systems cannot always be given a locally causal description. By now, it is well-known that entanglement is necessary, but may not always be sufficient to demonstrate this fact through a Bell inequality violation in a standard Bell experiment. A nonstandard Bell experiment, which involves local preprocessing and some kind of postselection, may however unveil the nonclassical correlations hidden in some entangled quantum states. Given this state of affairs, this thesis aims to clarify further the relationships between the notions of *correlations*, *Bell inequality violation* and *quantum entanglement* in discrete variable quantum systems. In this chapter, we will summarize our key findings and outline some possible avenues for future research.

Our study began in Chapter 5, where we looked into the problem of determining if a given quantum state ρ can violate some fixed but arbitrary Bell inequality in a standard Bell experiment. This is a high-dimensional variational problem where, in general, nontrivial optimization over the choice of local observables is required. To this end, we have derived two algorithms which can be used to determine, respectively, a lower bound (LB) and an upper bound (UB) on the *strength* of correlation that a quantum state ρ can offer in a given Bell experiment corresponding to some Bell inequality I_k — a quantity which we have given the notation $\mathcal{S}_{\text{QM}}^{(I_k)}(\rho)$. Both of these algorithms make use of convex optimization techniques in the form of a semidefinite program (SDP), which is readily solved on a computer. The LB algorithm requires one to solve a series of SDPs iteratively, whereas the UB algorithm provides a hierarchy of SDPs, with each giving a better upper bound on $\mathcal{S}_{\text{QM}}^{(I_k)}(\rho)$. These algorithms can also be implemented analytically. In fact, we have made use of the UB algorithm to derive a necessary condition for bipartite qudit states with vanishing coherence vectors to violate the Bell-CHSH inequality; a simple implementation of the LB algorithm has also enabled us to rederive the Horodecki criterion for two-qubit states. Since the bounds derived from these algorithms are usually not tight, these algorithms often need to be used

in tandem to determine if ρ can violate some Bell inequality I_k .

Next, in Chapter 6, we looked at some of the best known Bell inequalities violations by bipartite quantum states. In particular, using the LB algorithm derived in Chapter 5, we have obtained the local measurements giving the best known Bell-CH, and hence Bell-CHSH inequality violation for arbitrary pure two-qudit states. Then, by establishing a formal equivalence between the n -outcome CGLMP inequality and the I_{22nn} inequality, we have also obtained the best known I_{22dd} violation for the d -dimensional isotropic states $\rho_{I_d}(p)$. Together with the UB algorithm derived in Chapter 5, these best known violations were then used to show that for (arbitrary) bipartite pure two-qudit entangled state ρ , a better Bell-CH inequality violation can be obtained via collective measurements on $\rho^{\otimes N}$, i.e., N copies of ρ for $N > 2$. The same, however, cannot be said for mixed entangled states. In fact, we have strong numerical evidence suggesting that the maximal Bell-CH inequality violation for some entangled states may not depend on the number of copies N . Further numerical evidence even indicates that the set of mixed two-qubit states is dominated by those whose maximal Bell-CH inequality violation remains unchanged even when $N \geq 3$.

After that, in Chapter 7, we studied the possibility of deriving nonclassical correlations from all entangled states via a nonstandard Bell experiment. In other words, we wanted to know if it is actually possible to demonstrate some kind of observable nonlocality for all entangled states. To this end, we have explicitly characterized the set of bipartite quantum states which do not violate the Bell-CHSH inequality even after arbitrary local filtering operations — a set which we have given the notation $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Then, by considering a specific type of local filtering operation, we have (non-constructively) shown that for every bipartite entangled state σ , there exists an ancilla state $\rho \in \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ such that $\rho \otimes \sigma \notin \mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$. Interestingly, this means that even if both ρ and σ can be simulated, individually, by classical correlations in the most general single-copy nonstandard Bell-CHSH experiment, the combined state $\rho \otimes \sigma$ cannot be described by classical correlations in some single-copy nonstandard Bell-CHSH experiment. Consequently, we can now define a bipartite entangled state σ as precisely that which cannot be simulated by classical correlations when one consider all possible experiments that may be performed on σ in conjunction with non-Bell-CHSH-violating states.

Let us now make some remarks regarding future research. To begin with, we note that possible follow-up projects in relation to the work presented in each chapters have already been presented in some details at the end of the corresponding chapters. As such, we will not try to repeat all of them here, but to merely remind the readers of some of the key ones. Firstly, as one may have noticed, our analysis in this thesis has been carried out exclusively for discrete variable quantum systems and to a large extent, only for bipartite quantum systems. There are, of course, many interesting problems that are associated with Bell inequality violation in multipartite and continuous variable quantum systems. For example, it is still not known if all discrete multipartite pure entangled states can violate a Bell inequality in a standard Bell experiment. In continuous variable quantum systems, it is not even known if all bipartite pure entangled states can violate a Bell inequality. As a result, preliminary investigations on the adaptability of the tools that we have developed here to this latter scenario could be of some use.

Results that we have obtained in Chapter 5, as well as those presented in Ref. [17] have

indicated that upper bound techniques similar to those that we have developed in this thesis do allow us to investigate the extent to which quantum mechanics can violate a fixed but arbitrary Bell inequality. Further work on this is clearly desirable as it will help us to learn something about the extreme points of the set of quantum correlations. This is work in progress [197].

Given that we have only got a nonconstructive proof for the nonclassical correlations hidden in an arbitrary entangled state σ , it would be great if an explicit construction of the ancilla state ρ used in our protocol can be obtained. A general construction of the ancilla state may be formidable, but it would be helpful to at least solve this for some simple cases like Werner states, or more desirably, some bound entangled states. Finally, the arguably most important problem that is left opened from the present work is whether it is also possible to derive some observable, nonclassical correlations from *all* multipartite entangled quantum states, be it discrete or continuous. Any progress in this regard would certainly help us to improve our understanding of the quantum world, which is always full of surprises.

References

- [1] Y.-C. Liang and A. C. Doherty. *Bounds on quantum correlations in Bell-inequality experiments*. Physical Review A **75**(4), 042103 (2007).
- [2] Y.-C. Liang and A. C. Doherty. *Better Bell-inequality violation by collective measurements*. Physical Review A **73**(5), 052116 (2006).
- [3] Ll. Masanes, Y.-C. Liang, and A. C. Doherty. *All bipartite entangled states display some hidden nonlocality*. eprint quant-ph/0703268 (2007).
- [4] Y.-C. Liang, Ll. Masanes, and A. C. Doherty. *SLOCC convertibility between two-qubit states*. eprint quant-ph/0710.5350 (2007).
- [5] J. S. Bell. *On the Einstein-Podolsky-Rosen paradox*. Physics **1**(3), 195 (1964).
- [6] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. *Proposed experiment to test local hidden-variable theories*. Physical Review Letters **23**, 880 (1969).
- [7] J. F. Clauser and M. A. Horne. *Experimental consequences of objective local theories*. Physical Review D **10**(2), 526 (1974).
- [8] A. Peres. *All the Bell inequalities*. Foundations of Physics **29**(4), 589 (1999).
- [9] W. J. Munro, D. F. V. James, A. G. White, and P. G. Kwiat. *Maximizing the entanglement of two mixed qubits*. Physical Review A **64**(3), 030302 (2001).
- [10] T. C. Wei, K. Nemoto, P. M. Goldbart, P. G. Kwiat, W. J. Munro, and F. Verstraete. *Maximal entanglement versus entropy for mixed quantum states*. Physical Review A **67**(2), 022110 (2003).
- [11] L. Derkacz and L. Jakobczyk. *Clauser-Horne-Shimony-Holt violation and the entropy-concurrence plane*. Physical Review A **72**(4), 042321 (2005).
- [12] E. Santos. *Entropy inequalities and Bell inequalities for two-qubit systems*. Physical Review A **69**(2), 022305 (2004).
- [13] E. Santos. *Erratum: “Entropy inequalities and Bell inequalities for two-qubit systems” [Phys. Rev. A 69, 022305, (2004)]*. Physical Review A **70**(5), 059901 (2004).

- [14] T. Ito, H. Imai, and D. Avis. *Bell inequalities stronger than the Clauser-Horne-Shimony-Holt inequality for three-level isotropic states*. Physical Review A **74**(4), 042109 (2006).
- [15] K. G. H. Vollbrecht and R. F. Werner. *Entanglement measures under symmetry*. Physical Review A **64**(6), 062307 (2001).
- [16] A. Acín, T. Durt, N. Gisin, and J. I. Latorre. *Quantum nonlocality in two three-level systems*. Physical Review A **65**(5), 052325 (2002).
- [17] M. Navascués, S. Pironio, and A. Acín. *Bounding the set of quantum correlations*. Physical Review Letters **98**(1), 010401 (2007).
- [18] E. Schrödinger. *Discussion of probability relations between separated systems*. Proceedings of the Cambridge Philosophical Society **31**, 555 (1935).
- [19] E. Schrödinger. *Die gegenwärtige situation in der quantenmechanik*. Naturwissenschaften **23**, 807 (1935).
- [20] A. Einstein, B. Podolsky, and N. Rosen. *Can quantum-mechanical description of physical reality be considered complete?* Physical Review **47**, 777 (1935).
- [21] J. D. Trimmer. *The present situation in quantum mechanics: A translation of Schrödinger's "cat paradox" paper*. Proceedings of the American Physical Society **124**(5), 323 (1980).
- [22] A. Einstein. *Albert Einstein: Philosopher-Scientist*, chap. Reply to Criticisms (Cambridge University Press, 1949).
- [23] J. S. Bell. *Einstein-Podolsky-Rosen experiments*. In *Proceedings of the Symposium on Frontier Problems in High Energy Physics*, pp. 33–45 (1976).
- [24] J. S. Bell. *On the problem of hidden variables in quantum mechanics*. Review of Modern Physics **38**, 447 (1966).
- [25] D. Bohm. *Quantum Theory* (Englewood Cliffs, N.J. : Prentice-Hall, 1951).
- [26] A. Aspect. *Bell's inequality test: more ideal than ever*. Nature **398**, 189 (1999).
- [27] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. A. Itano, C. Monroe, and D. J. Wineland. *Experimental violation of a Bell's inequality with efficient detection*. Nature **409**(6822), 791 (2001).
- [28] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin. *Experimental entanglement distillation and 'hidden' non-locality*. Nature **409**(6823), 1014 (2001).
- [29] R. F. Werner. *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*. Physical Review A **40**, 4277 (1989).

-
- [30] R. Horodecki, P. Horodecki, and M. Horodecki. *Violating Bell inequality by mixed spin- $\frac{1}{2}$ states: Necessary and sufficient condition*. Physics Letters A **200**(5), 340 (1995).
- [31] J. Barrett. *Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality*. Physical Review A **65**(4), 042302 (2002).
- [32] G. Tóth and A. Acín. *Genuine tripartite entangled states with a local hidden-variable model*. Physical Review A **74**(3), 030306 (2006).
- [33] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín. *Noise robustness of the nonlocality of entangled quantum states*. Physical Review Letters **99**(4), 040403 (2007).
- [34] V. Capasso, D. Fortunato, and F. Selleri. *Sensitive observables of quantum mechanics*. International Journal of Theoretical Physics **7**(5), 319 (1973).
- [35] N. Gisin. *Bell inequality holds for all non-product states*. Physics Letters A **154**(5-6), 201 (1991).
- [36] N. Gisin and A. Peres. *Maximal violation of Bell inequality for arbitrarily large spin*. Physics Letters A **162**(1), 15 (1992).
- [37] J.-L. Chen, C. Wu, L. C. Kwek, and C. H. Oh. *Gisin's theorem for three qubits*. Physical Review Letters **93**(14), 140407 (2004).
- [38] S. Popescu. *Bell's inequalities and density matrices: Revealing "hidden" nonlocality*. Physical Review Letters **74**(14), 2619 (1995).
- [39] N. Gisin. *Hidden quantum nonlocality revealed by local filters*. Physics Letters A **210**(3), 151 (1996).
- [40] A. Peres. *Collective tests for quantum nonlocality*. Physical Review A **54**(4), 2685 (1996).
- [41] A. K. Ekert. *Quantum cryptography based on Bell's theorem*. Physical Review Letters **67**(6), 661 (1991).
- [42] V. Scarani and N. Gisin. *Quantum communication between n partners and Bell's inequalities*. Physical Review Letters **87**(11), 117901 (2001).
- [43] A. Acín, N. Gisin, Ll. Masanes, and V. Scarani. *Bell's inequalities detect efficient entanglement*. International Journal of Quantum Information **2**(1), 23 (2004).
- [44] J. Barrett, L. Hardy, and A. Kent. *No signaling and quantum key distribution*. Physical Review Letters **95**(1), 010503 (2005).
- [45] A. Acín, N. Gisin, and Ll. Masanes. *From Bell's theorem to secure quantum key distribution*. Physical Review Letters **97**(12), 120405 (2006).
- [46] Ll. Masanes and A. Winter. *Unconditional security of key distribution from causality constraints*. eprint quant-ph/0606049 (2006).

-
- [47] H. Buhrman, R. Cleve, and W. van Dam. *Quantum entanglement and communication complexity*. SIAM Journal On Computing **30**(6), 1829 (2000).
- [48] Č. Brukner, M. Żukowski, and A. Zeilinger. *Quantum communication complexity protocol with two entangled qutrits*. Physical Review Letters **89**(19), 197901 (2002).
- [49] Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger. *Bell's inequalities and quantum communication complexity*. Physical Review Letters **92**(12), 127901 (2004).
- [50] C. H. Bennett, G. Brassard, C. Crpeau, R. Jozsa, A. Peres, and W. K. Wootters. *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*. Physical Review Letters **70**(13), 1895 (1993).
- [51] R. Horodecki, M. Horodecki, and P. Horodecki. *Teleportation, Bell's inequalities and inseparability*. Physics Letters A **222**, 21 (1996).
- [52] S. Popescu. *Bell's inequalities versus teleportation: What is nonlocality?* Physical Review Letters **72**(6), 797 (1994).
- [53] N. Gisin. *Bell inequalities: many questions, a few answers*. eprint quant-ph/0702021 (2007).
- [54] S. Filipp and K. Svozil. *Generalizing Tsirelson's bound on Bell inequalities using a min-max principle*. Physical Review Letters **93**, 130407 (2004).
- [55] H. Buhrman and S. Massar. *Causality and Tsirelson's bounds*. Physical Review A **72**(5), 052103 (2005).
- [56] S. Wehner. *Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities*. Physical Review A **73**(2), 022110 (2006).
- [57] B. F. Toner. *Monogamy of nonlocal quantum correlations*. eprint quant-ph/0601172 (2006).
- [58] D. Avis, H. Imai, and T. Ito. *On the relationship between convex bodies related to correlation experiments with dichotomic observables*. Journal of Physics A-Mathematical and General **39**(36), 11283 (2006).
- [59] Ll. Masanes. *Tight Bell inequality for d-outcome measurements correlations*. Quantum Information & Computation **3**(4), 345 (2003).
- [60] D. Collins and N. Gisin. *A relevant two qubit Bell inequality inequivalent to the CHSH inequality*. Journal of Physics A-Mathematical and General **37**(5), 1775 (2004).
- [61] D. Avis, H. Imai, T. Ito, and Y. Sasaki. *Two-party Bell inequalities derived from combinatorics via triangular elimination*. Journal of Physics A-Mathematical and General **38**, 10971 (2005).
- [62] I. Pitowsky. *Quantum Probability – Quantum Logic*, vol. 321 of *Lecture Notes in Physics* (Springer, Berlin, 1989).

-
- [63] S. Boyd and L. Vandenberghe. *Convex Optimization* (New York : Cambridge, 2004).
- [64] L. Vandenberghe and S. Boyd. *Semidefinite programming*. SIAM Review **38**, 49 (1996).
- [65] A. Peres. *Quantum Theory : Concepts and Methods* (Springer, New York, 1995).
- [66] C. J. Isham. *Lectures on Quantum Theory: Mathematical and Structural Foundations* (Imperial College Press, London, 1995).
- [67] N. Bohr. *Quantum postulate and recent developments in atomism*. Naturwissenschaften **16**, 245 (1928).
- [68] N. Bohr. *The quantum postulate and the recent development of atomic theory*. Nature **121**, 580 (1928).
- [69] M. O. Scully, B.-G. Englert, and H. Walther. *Quantum optical tests of complementarity*. Nature **351**(6322), 111 (1991).
- [70] L. de Broglie. *La mcanique ondulatoire et la structure atomique de la matire et du rayonnement*. Journal de Physique et du Radium **8**, 225 (1927).
- [71] L. de Broglie. *Rapports et Discussions du Cinquime Conseil de Physique Solvay*, chap. La nouvelle dynamique des quanta, pp. 105–141 (Paris: Gauthier-Villars, 1928).
- [72] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer Verlag, 1932).
- [73] D. Bohm. *A suggested interpretation of the quantum theory in terms of "hidden variables" I*. Physical Review **85**, 166 (1952).
- [74] D. Bohm. *A suggested interpretation of the quantum theory in terms of "hidden variables" II*. Physical Review **85**, 180 (1952).
- [75] J. S. Bell. *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press, 1987).
- [76] J. S. Bell. *Foundation of Quantum Mechanics. Proceedings of the International School of Physics 'Enrico Fermi', course II*, chap. Introduction to the Hidden-Variable Question, pp. 171–181 (Academic, New York, 1971).
- [77] J. S. Bell. *Between Science and Technology*, chap. La Nouvelle Cuisine, pp. 97–115 (Elsevier/North-Holland, 1990).
- [78] T. Norsen. *Against 'realism'*. Foundations of Physics **37**(3), 311 (2007).
- [79] J. S. Bell. *EPR correlations and EPW distributions*. Annals of the New York Academy of Sciences **480**, 263 (1986).
- [80] B. M. Terhal. *Bell inequalities and the separability criterion*. Physics Letters A **271**(5-6), 319 (2000).

- [81] P. Hyllus, O. Gühne, D. Bruß, and M. Lewenstein. *Relations between entanglement witnesses and Bell inequalities*. Physical Review A **72**(1), 012321 (2005).
- [82] J. F. Clauser and A. Shimony. *Bell's theorem: experimental tests and implications*. Reports on Progress in Physics **41**, 1881 (1978).
- [83] S. J. Freedman and J. F. Clauser. *Experimental test of local hidden-variable theories*. Physical Review Letters **28**(14), 938 (1972).
- [84] A. Shimony. *Bell's theorem*. In E. N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Summer 2005). URL <http://plato.stanford.edu/archives/sum2005/entries/bell-theorem/>.
- [85] A. Aspect, J. Dalibard, and G. Roger. *Experimental test of Bell's inequalities using time-varying analyzers*. Physical Review Letters **49**(25), 1804 (1982).
- [86] T. W. Marshall, E. Santos, and F. Selleri. *Local realism has not been refuted by atomic cascade experiments*. Physics Letters A **98**(1-2), 5 (1983).
- [87] E. Santos. *Does quantum mechanics violate the Bell inequalities?* Physical Review Letters **66**(11), 1388 (1991).
- [88] E. Santos. *Critical analysis of the empirical tests of local hidden-variable theories*. Physical Review A **46**(7), 3646 (1992).
- [89] E. Santos. *The failure to perform a loophole-free test of Bell's inequality supports local realism*. Foundations of Physics **34**(11), 1643 (2004).
- [90] E. Santos. *Bell's theorem and the experiments: Increasing empirical support for local realism?* Studies In History and Philosophy of Modern Physics **36B**(3), 544 (2005).
- [91] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger. *Violation of Bell's inequality under strict Einstein locality conditions*. Physical Review Letters **81**(23), 5039 (1998).
- [92] E. S. Fry and T. Walther. *Quantum [Un]speakables*, chap. Atom Based Tests of the Bell Inequalities - the Legacy of John Bell Continues ..., pp. 103–117 (Springer: Berlin-Heidelberg-New York, 2002).
- [93] J. Barrett, D. Collins, L. Hardy, A. Kent, and S. Popescu. *Quantum nonlocality, Bell inequalities, and the memory loophole*. Physical Review A **66**(4), 042111 (2002).
- [94] A. Kent. *Causal quantum theory and the collapse locality loophole*. Physical Review A **72**(1), 012107 (2005).
- [95] M. Genovese. *Research on hidden variable theories: A review of recent progresses*. Physics Reports-Review Section of Physics Letters **413**(6), 319 (2005).

-
- [96] M. Froissart. *Constructive generalization of Bell's inequalities*. Nuovo Cimento Della Società Italiana Di Fisica B-General Physics, Relativity, Astronomy and Mathematical Physics and Methods **64**(2), 241 (1981).
- [97] A. Garg and N. D. Mermin. *Farkas lemma and the nature of reality: Statistical implications of quantum correlations*. Foundations of Physics **14**(1), 1 (1984).
- [98] I. Pitowsky. *Correlation polytopes: their geometry and complexity*. Mathematical Programming **50**(3), 395 (1991).
- [99] A. Fine. *Hidden variables, joint probability, and the Bell inequalities*. Physical Review Letters **48**(5), 291 (1982).
- [100] A. Garg and N. D. Mermin. *Comment on "hidden variables, joint probability, and the Bell inequalities"*. Physical Review Letters **49**(3), 242 (1982).
- [101] A. Fine. *Fine responds*. Physical Review Letters **49**(3), 243 (1982).
- [102] A. Garg and N. D. Mermin. *Local realism and measured correlations in the spin-s Einstein-Podolsky-Rosen experiment*. Physical Review D **27**(2), 339 (1983).
- [103] S. Popescu and D. Rohrlich. *Quantum nonlocality as an axiom*. Foundations of Physics **24**(3), 379 (1994).
- [104] R. F. Werner and M. M. Wolf. *Bell inequalities and entanglement*. Quantum Information and Computation **1**, 1 (2001).
- [105] V. Scarani and N. Gisin. *Superluminal influences, hidden variables, and signaling*. Physics Letters A **295**(4), 167 (2002).
- [106] B. Grünbaum. *Convex Polytope* (Springer, New York, 2003), 2 ed.
- [107] G. M. Ziegler. *Lectures on Polytopes* (Springer, New York, 1995), 1 ed.
- [108] B. M. Terhal, A. C. Doherty, and D. Schwab. *Symmetric extensions of quantum states and local hidden variable theories*. Physical Review Letters **90**(15), 157903 (2003).
- [109] I. Pitowsky and K. Svozil. *Optimal tests of quantum nonlocality*. Physical Review A **64**(1), 014102 (2001).
- [110] N. D. Mermin. *Quantum mechanics vs local realism near the classical limit: A Bell inequality of spin s*. Physical Review D **22**(2), 356 (1980).
- [111] A. Garg and N. D. Mermin. *Bell inequalities with a range of violation that does not diminish as the spin becomes arbitrarily large*. Physical Review Letters **49**(13), 901 (1982).
- [112] A. Garg and N. D. Mermin. *Correlation inequalities and hidden variables*. Physical Review Letters **49**(17), 1220 (1982).

- [113] N. D. Mermin and G. M. Schwarz. *Joint distributions and local realism in the higher-spin Einstein-Podolsky-Rosen experiment*. Foundations of Physics **12**(2), 101 (1982).
- [114] N. D. Mermin. *Generalizations of Bell's Theorem to higher spins and higher correlations*, chap. Fundamental Questions in Quantum Mechanics, pp. 7–20 (Gordon and Breach, 1986).
- [115] R. F. Werner and M. M. Wolf. *All-multipartite Bell-correlation inequalities for two dichotomic observables per site*. Physical Review A **64**(3), 032112 (2001).
- [116] C. Śliwa. *Symmetries of the Bell correlation inequalities*. Physics Letters A **317**(3-4), 165 (2003).
- [117] S. Pironio. *Lifting Bell inequalities*. Journal of Mathematical Physics **46**(6), 062112 (2005).
- [118] N. Brunner and N. Gisin. *Partial list of bipartite Bell inequalities with four binary settings*. eprint arXiv:0711.3362 (2007).
- [119] D. Avis and T. Ito. *New classes of facets of the cut polytope and tightness of I_{mm22} Bell inequalities*. Discrete Applied Mathematics **155**(13), 1689 (2007).
- [120] <http://www-imai.is.s.u-tokyo.ac.jp/~tsuyoshi/bell/bell15.html>.
- [121] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. *Bell inequalities for arbitrarily high-dimensional systems*. Physical Review Letters **88**(4), 040404 (2002).
- [122] D. Kaszlikowski, L. C. Kwek, J.-L. Chen, M. Żukowski, and C. H. Oh. *Clauser-Horne inequality for three-state systems*. Physical Review A **65**(3), 032118 (2002).
- [123] N. D. Mermin. *Extreme quantum entanglement in a superposition of macroscopically distinct states*. Physical Review Letters **65**(15), 1838 (1990).
- [124] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pp. 69–72 (Kluwer Academics, Dordrecht, 1989).
- [125] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Going beyond Bell's theorem*. eprint arXiv:0712.0921 (2007).
- [126] S. M. Roy and V. Singh. *Tests of signal locality and Einstein-Bell locality for multi-particle systems*. Physical Review Letters **67**(20), 2761 (1991).
- [127] M. Ardehali. *Bell inequalities with a magnitude of violation that grows exponentially with the number of particles*. Physical Review A **46**(9), 5375 (1992).
- [128] A. V. Belinskii and D. N. Klyshko. *The interference of light and the Bell theorem*. Uspekhi Fizicheskikh Nauk **163**(8), 1 (1993).
- [129] A. V. Belinskii and D. N. Klyshko. *Interference of light and Bell's theorem*. Physics-Uspekhi **36**(8), 653 (1993).

-
- [130] N. Gisin and H. Bechmann-Pasquinucci. *Bell inequality, Bell states and maximally entangled states for n qubits*. Physics Letters A **246**(1-2), 1 (1998).
- [131] M. Żukowski and Č. Brukner. *Bell's theorem for general n -qubit states*. Physical Review Letters **88**(21), 210401 (2002).
- [132] H. Weinfurter and M. Żukowski. *Four-photon entanglement from down-conversion*. Physical Review A **64**01(1), 010102 (2001).
- [133] X. H. Wu and H. S. Zong. *Violation of local realism by a system with n spin-1/2 particles*. Physical Review A **68**(3), 032102 (2003).
- [134] W. Laskowski, T. Paterek, M. Żukowski, and Č. Brukner. *Tight multipartite Bell's inequalities involving many measurement settings*. Physical Review Letters **93**(20), 200401 (2004).
- [135] M. Żukowski. *On tight multipartite Bell inequalities for many settings*. Quantum Information Processing **5**(4), 287 (2006).
- [136] M. Żukowski. *All tight multipartite Bell correlation inequalities for three dichotomic observables per observer*. eprint quant-ph/0611086 (2006).
- [137] M. Wieśniak, P. Badziąg, and M. Żukowski. *Explicit form of correlation-function three-setting tight Bell inequalities for three qubits*. Physical Review A **76**(1), 012110 (2007).
- [138] A. Acín, J.-L. Chen, N. Gisin, D. Kaszlikowski, L. C. Kwek, C. H. Oh, and M. Żukowski. *Coincidence Bell inequality for three three-dimensional systems*. Physical Review Letters **92**(25), 250404 (2004).
- [139] V. Scarani, A. Acín, E. Schenck, and M. Aspelmeyer. *Nonlocality of cluster states of qubits*. Physical Review A **71**(4), 042325 (2005).
- [140] K. Chen, S. Alberverio, and S. M. Fei. *Two-setting Bell inequalities for many qubits*. Physical Review A **74**(5), 050101 (2006).
- [141] C. Wu, Y. Yeo, L. C. Kwek, and C. H. Oh. *Quantum nonlocality of four-qubit entangled states*. Physical Review A **75**(3), 032332 (2007).
- [142] L. A. Khalfin and B. S. Tsirelson. *Quantum classical correspondence in the light of Bells inequalities*. Foundations of Physics **22**(7), 879 (1992).
- [143] B. S. Cirelson. *Quantum generalizations of Bell's inequality*. Lett. Math. Phys. **4**, 93 (1980).
- [144] B. S. Tsirelson. *Quantum analogues of Bell's inequalities. The case of two spatially divided domains*. Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR **142**, 174 (1985).

- [145] B. S. Tsirel'son. *Quantum analogues of the Bell inequalities. The case of two spatially separated domains*. Journal of Soviet Mathematics **36**(4), 557 (1987).
- [146] I. Pitowsky. *The range of quantum probability*. Journal of Mathematical Physics **27**(6), 1556 (1986).
- [147] M. Żukowski, R. Horodecki, M. Horodecki, and P. Horodecki. *Generalized quantum measurements and local realism*. Physical Review A **58**(3), 1694 (1998).
- [148] A. Cabello. *How much larger quantum correlations are than classical ones*. Physical Review A **72**(1), 012113 (2005).
- [149] M. Żukowski. *Some news about Bell inequalities*. Journal of Modern Optics **50**(6-7), 1151 (2003).
- [150] D. Jonathan and M. B. Plenio. *Minimal conditions for local pure-state entanglement manipulation*. Physical Review Letters **83**(7), 1455 (1999).
- [151] J. Degorre, S. Laplante, and J. Roland. *Simulating quantum correlations as a distributed sampling problem*. Physical Review A **72**(6), 062314 (2005).
- [152] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information* (Cambridge ; New York : Cambridge University Press, 2000).
- [153] A. Acín, N. Gisin, and B. Toner. *Grothendieck's constant and local models for noisy entangled quantum states*. Physical Review A **73**(6), 062105 (2006).
- [154] J. L. Krivine. *Grothendieck constants and positive type functions on spheres*. Advances In Mathematics **31**(1), 16 (1979).
- [155] M. Horodecki and P. Horodecki. *Reduction criterion of separability and limits for a class of distillation protocols*. Physical Review A **59**(6), 4206 (1999).
- [156] V. M. Kendon, K. Życzkowski, and W. J. Munro. *Bounds on entanglement in qudit subsystems*. Physical Review A **66**(6), 062310 (2002).
- [157] M. A. Nielsen. *Conditions for a class of entanglement transformations*. Physical Review Letters **83**(2), 436 (1999).
- [158] T. Eggeling and R. F. Werner. *Separability properties of tripartite states with $U \otimes U \otimes U$ symmetry*. Physical Review A **63**(4), 042111 (2001).
- [159] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera. *Classification of mixed three-qubit states*. Physical Review Letters **87**(4), 040401 (2001).
- [160] P. Horodecki. *Separability criterion and inseparable mixed states with positive partial transposition*. Physics Letters A **232**(5), 333 (1997).
- [161] A. Peres. *Separability criterion for density matrices*. Physical Review Letters **77**(8), 1413 (1996).

-
- [162] M. Horodecki, P. Horodecki, and R. Horodecki. *Separability of mixed states: Necessary and sufficient conditions*. Physics Letters A **223**(1-2), 1 (1996).
- [163] M. Horodecki, P. Horodecki, and R. Horodecki. *Mixed state entanglement and distillation: Is there a "bound" entanglement in nature?* Physical Review Letters **80**(24), 5239 (1998).
- [164] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. *Purification of noisy entanglement and faithful teleportation via noisy channels*. Physical Review Letters **76**(5), 722 (1996).
- [165] R. F. Werner and M. M. Wolf. *Bell's inequalities for states with positive partial transpose*. Physical Review A **61**(6), 062102 (2000).
- [166] W. Dür. *Multipartite bound entangled states that violate Bell's inequality*. Physical Review Letters **87**(23), 230402 (2001).
- [167] D. M. Greenberger, M. Horne, A. Shimony, and A. Zeilinger. *Bell theorem without inequalities*. American Journal of Physics **58**(12), 1131 (1990).
- [168] A. Acín. *Distillability, Bell inequalities, and multiparticle bound entanglement*. Physical Review Letters **88**(2), 027901 (2002).
- [169] A. Acín, V. Scarani, and M. M. Wolf. *Bell's inequalities and distillability in n -qubit systems*. Physical Review A **66**(4), 042323 (2002).
- [170] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal. *Unextendible product bases and bound entanglement*. Physical Review Letters **82**(26), 5385 (1999).
- [171] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. *Distinguishing separable and entangled states*. Physical Review Letters **88**(18), 187904 (2002).
- [172] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. *Complete family of separability criteria*. Physical Review A **69**(2), 022308 (2004).
- [173] B. M. Terhal, A. C. Doherty, and D. Schwab. *Local hidden variable theories for quantum states*. eprint quant-ph/0210053 (2002).
- [174] P. Horodecki, M. Horodecki, and R. Horodecki. *Bound entanglement can be activated*. Physical Review Letters **82**(5), 1056 (1999).
- [175] V. Scarani and N. Gisin. *Spectral decomposition of Bell's operators for qubits*. Journal of Physics A-Mathematical and General **34**(30), 6043 (2001).
- [176] M. Żukowski, Č. Brukner, W. Laskowski, and M. Wiesniak. *Do all pure entangled states violate Bell's inequalities for correlation functions?* Physical Review Letters **88**(21), 210402 (2002).

-
- [177] C. Wu, J.-L. Chen, L. C. Kwek, and C. H. Oh. *Quantum nonlocality of n -qubit W states*. Physical Review A **73**(1), 012310 (2006).
- [178] S. L. Braunstein, A. Mann, and M. Revzen. *Maximal violation of Bell inequalities for mixed states*. Physical Review Letters **68**(22), 3259 (1992).
- [179] M. M. Deza and M. Laurent. *Geometry of cuts and metrics*. In *Algorithms and Combinatorics*, vol. 15 (Springer, New York, 1997).
- [180] S. Popescu and D. Rohrlich. *Which states violate Bell's inequality maximally?* Physics Letters A **169**(6), 411 (1992).
- [181] A. Cabello. *Proposed experiment to test the bounds of quantum correlations*. Physical Review Letters **92**(6), 060403 (2004).
- [182] F. A. Bovino, G. Castagnoli, I. P. Degiovanni, and S. Castelletto. *Experimental evidence for bounds on quantum correlations*. Physical Review Letters **92**(6), 060404 (2004).
- [183] P. A. Parrilo. *Semidefinite programming relaxations for semialgebraic problems*. Mathematical Programming **96**(2), 293 (2003).
- [184] J. B. Lasserre. *Global optimization with polynomials and the problem of moments*. Siam J. Optim. **11**(3), 796 (2001).
- [185] J. Eisert, P. Hyllus, O. Gühne, and M. Curty. *Complete hierarchies of efficient approximations to problems in entanglement theory*. Physical Review A **70**(6), 062317 (2004).
- [186] R. A. Horn and C. R. Johnson. *Matrix Analysis* (Cambridge University Press, New York, NY, USA, 1990).
- [187] G. Strang. *Linear Algebra and Its Applications* (Harcourt Brace, San Diego, 1988), 3rd ed.
- [188] C. W. Helstrom. *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [189] R. Cleve, P. Høyer, B. Toner, and J. Watrous. *Consequences and limits of nonlocal strategies*. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity*, pp. 236–249 (2004).
- [190] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB* (2004).
- [191] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. *SOSTOOLS and its control applications*. Lecture Notes in Control and Information Sciences **312**, 273 (2005).

-
- [192] K. Schmüdgen. *The k -moment problem for compact semi-algebraic sets*. Mathematische Annalen **289**(2), 203 (1991).
- [193] M. S. Byrd and N. Khaneja. *Characterization of the positivity of the density matrix in terms of the coherence vector representation*. Physical Review A **68**(6), 062322 (2003).
- [194] G. Kimura. *The Bloch vector for n -level systems*. Physics Letters A **314**(5-6), 339 (2003).
- [195] F. Verstraete and M. M. Wolf. *Entanglement versus Bell violations and their behavior under local filtering operations*. Physical Review Letters **89**(17), 170401 (2002).
- [196] J. W. Nie, J. Demmel, and B. Sturmfels. *Minimizing polynomials via sum of squares over the gradient ideal*. Mathematical Programming **106**(3), 587 (2006).
- [197] A. C. Doherty, Y.-C. Liang, B. Toner, and S. Wehner. *The quantum moment problem and bounds on entangled multi-prover games*. (in preparation) (2007).
- [198] A. Peres. *Finite violation of a Bell inequality for arbitrarily large spin*. Physical Review A **46**, 4413 (1992).
- [199] D. Kaszlikowski, P. Gnaniński, M. Żukowski, W. Miklaszewski, and A. Zeilinger. *Violations of local realism by two entangled n -dimensional systems are stronger than for two qubits*. Physical Review Letters **85**(21), 4418 (2000).
- [200] L.-B. Fu, J.-L. Chen, and X.-G. Zhao. *Maximal violation of the Clauser-Horne-Shimony-Holt inequality for two qutrits*. Physical Review A **68**(2), 022323 (2003).
- [201] L.-B. Fu, J.-L. Chen, and S.-G. Chen. *Maximal violation of Clauser-Horne-Shimony-Holt inequality for four-level systems*. Physical Review A **69**(3), 034305 (2004).
- [202] S. Popescu and D. Rohrlich. *Generic quantum nonlocality*. Physics Letters A **166**(5-6), 293 (1992).
- [203] E. Schmidt. *Zur theorie der linearen und nichtlinearen integralgleichungen*. Math. Annalen. **63**, 433 (1906).
- [204] S. M. Barnett and S. J. D. Phoenix. *Bell inequality and the Schmidt decomposition*. Physics Letters A **167**(3), 233 (1992).
- [205] J.-L. Chen, D. Kaszlikowski, L. C. Kwek, C. H. Oh, and M. Żukowski. *Entangled three-state systems violate local realism more strongly than qubits: An analytical proof*. Physical Review A **64**(5), 052109 (2001).
- [206] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. *Concentrating partial entanglement by local operations*. Physical Review A **53**(4), 2046 (1996).
- [207] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. *Volume of the set of separable states*. Physical Review A **58**(2), 883 (1998).

- [208] O. Krüger and R. Werner. *Some open problems in quantum information theory*. eprint quant-ph/0504166 (2005).
- [209] Ll. Masanes. *Asymptotic violation of Bell inequalities and distillability*. Physical Review Letters **97**(5), 050503 (2006).
- [210] Ll. Masanes. *All bipartite entangled states are useful for information processing*. Physical Review Letters **96**(15), 150501 (2006).
- [211] K. Kraus. *General state changes in quantum theory*. Annals of Physics **64**(2), 311 (1971).
- [212] K. Kraus. *States, effects, and operations : fundamental notions of quantum theory : lectures in mathematical physics at the University of Texas at Austin* (Springer, Berlin, 1983).
- [213] W. Dür, G. Vidal, and J. I. Cirac. *Three qubits can be entangled in two inequivalent ways*. Physical Review A **62**(6), 062314 (2000).
- [214] M. Horodecki, P. Horodecki, and R. Horodecki. *Inseparable two spin-1/2 density matrices can be distilled to a singlet form*. Physical Review Letters **78**(4), 574 (1997).
- [215] B. G. Englert and N. Metwally. *Separability of entangled q-bit pairs*. Journal of Modern Optics **47**(12), 2221 (2000).
- [216] R. Bhatia. *Matrix Analysis* (Springer, New York, 1996).
- [217] R. Horodecki and M. Horodecki. *Information-theoretic aspects of inseparability of mixed states*. Physical Review A **54**(3), 1838 (1996).
- [218] B. D. Craven and J. J. Koliha. *Generalizations of Farkas theorem*. Siam Journal on Mathematical Analysis **8**(6), 983 (1977).
- [219] B. F. Toner and D. Bacon. *Communication cost of simulating Bell correlations*. Physical Review Letters **91**(18), 187904 (2003).
- [220] <http://www.physics.uq.edu.au/qisci/yerng/>.
- [221] A. Jamiołkowski. *Linear transformations which preserve trace and positive semidefiniteness of operators*. Reports on Mathematical Physics **3**(4), 275 (1972).
- [222] M.-D. Choi. *Completely positive linear maps on complex matrices*. Linear Algebra and its Applications **10**(3), 285 (1975).
- [223] V. P. Belavkin and P. Staszewski. *A Radon-Nikodym theorem for completely positive maps*. Reports on Mathematical Physics **24**(1), 49 (1986).
- [224] E. M. Rains. *Entanglement purification via separable superoperators*. e-print quant-ph/9707002 (1997).

-
- [225] V. Vedral and M. B. Plenio. *Entanglement measures and purification procedures*. Physical Review A **57**(3), 1619 (1998).
 - [226] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters. *Quantum nonlocality without entanglement*. Physical Review A **59**(2), 1070 (1999).
 - [227] J. I. Cirac, W. Dur, B. Kraus, and M. Lewenstein. *Entangling operations and their implementation using a small amount of entanglement*. Physical Review Letters **86**(3), 544 (2001).



Bell-diagonal Preserving Separable Maps

In this Appendix we classify the four-qubit states that commute with $U \otimes U \otimes V \otimes V$, where U and V are arbitrary members of the Pauli group. We characterize the set of separable states for this class, in terms of a finite number of entanglement witnesses. Equivalently, we characterize the set of two-qubit, Bell-diagonal-preserving, completely positive maps (henceforth abbreviated as CPM) that are separable. These separable CPMs correspond to protocols that can be implemented with stochastic local quantum operations assisted by classical communication (SLOCC). Explicit characterization of these CPMs is an essential ingredient of the proof of Lemma 17.

A.1 Four-qubit Separable States with $U \otimes U \otimes V \otimes V$ Symmetry

In this section, we will characterize the set of separable states commuting with $U \otimes U \otimes V \otimes V$, where U and V are arbitrary members of the Pauli group. Let us begin by reminding the reader about an important property of two-qubit states which commute with all unitaries of the form $U \otimes U$, where U is an arbitrary member of the Pauli group. The Pauli group is generated by the Pauli matrices $\{\sigma_i\}_{i=x,y,z}$, Eq. (2.6), and has 16 elements. The representation $U \otimes U$ comprises four 1-dimensional irreducible representations, each acting on the subspace spanned by one vector of the Bell basis¹

$$|\Phi_1\rangle \equiv \frac{1}{\sqrt{2}} (|0\rangle|0\rangle \pm |1\rangle|1\rangle), \quad (\text{A.1a})$$

$$|\Phi_3\rangle \equiv \frac{1}{\sqrt{2}} (|0\rangle|1\rangle \pm |1\rangle|0\rangle). \quad (\text{A.1b})$$

¹These states are more conventionally denoted by $|\Phi_1\rangle = |\Phi^+\rangle$, $|\Phi_2\rangle = |\Phi^-\rangle$, $|\Phi_3\rangle = |\Psi^+\rangle$, $|\Phi_4\rangle = |\Psi^-\rangle$.

This implies that [15] any two-qubit state which commutes with $U \otimes U$ can be written as $\rho = \sum_{k=1}^4 [r]_k \Pi_k$, where $\Pi_k \equiv |\Phi_k\rangle\langle\Phi_k|$ is the k^{th} Bell projector. With this information in mind, we are now ready to discuss the case that is of our interest.

We would like to characterize the set of four-qubit states which commute with all unitaries $U \otimes U \otimes V \otimes V$, where U and V are members of the Pauli group. Let us denote this set of states by \mathcal{R} and the state space of $\rho \in \mathcal{R}$ as $\mathcal{H} \simeq \mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{A}''} \otimes \mathcal{H}_{\mathcal{B}''}$, where $\mathcal{H}_{\mathcal{A}'}$, $\mathcal{H}_{\mathcal{B}'}$ etc. are Hilbert spaces of the constituent qubits. In this notation, both the subsystems associated with $\mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{B}'}$ and that with $\mathcal{H}_{\mathcal{A}''} \otimes \mathcal{H}_{\mathcal{B}''}$ have $U \otimes U$ symmetry and hence are linear combinations of Bell-diagonal projectors [15].

Our aim in this section is to provide a full characterization of the set of ρ that are separable between $\mathcal{H}_{\mathcal{A}} \equiv \mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{A}''}$ and $\mathcal{H}_{\mathcal{B}} \equiv \mathcal{H}_{\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{B}''}$ (see Figure A.1). Throughout this section, a state is said to be *separable* if and only if it is separable between $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$.

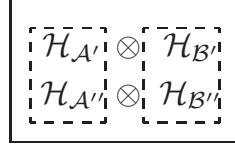


FIGURE A.1: A schematic diagram for the subsystems constituting ρ . Subsystems that are arranged in the same row in the diagram have $U \otimes U$ symmetry and hence are represented by Bell-diagonal states [15] (see text for details). In this Appendix, we are interested in states that are separable between subsystems enclosed in the two dashed boxes.

The symmetry of ρ allows one to write it as a *non-negative* combination of (tensored-) Bell projectors:

$$\rho = \sum_{i=1}^4 \sum_{j=1}^4 [r]_{i,j} \Pi_i \otimes \Pi_j, \quad (\text{A.2})$$

where the Bell projector before and after the tensor product, respectively, acts on $\mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{B}'}$ and $\mathcal{H}_{\mathcal{A}''} \otimes \mathcal{H}_{\mathcal{B}''}$ (Figure A.1). Thus, any state $\rho \in \mathcal{R}$ can be represented in a compact manner, via the corresponding 4×4 matrix r . More generally, any operator μ acting on the Hilbert space \mathcal{H} and having the $U \otimes U \otimes V \otimes V$ symmetry admits a 4×4 matrix representation M via:

$$\mu = \sum_{i=1}^4 \sum_{j=1}^4 [M]_{i,j} \Pi_i \otimes \Pi_j, \quad (\text{A.3})$$

where $[M]_{i,j}$ is now not necessarily non-negative. When there is no risk of confusion, we will also refer to r and M , respectively, as a state and an operator having this symmetry.

Evidently, in this representation, an operator μ is non-negative if and only if all entries in the corresponding 4×4 matrix M are non-negative. Notice also that by appropriate local unitary transformation, one can swap any Π_i with any other Π_j , $j \neq i$ while keeping all the other Π_k , $k \neq i, j$ unaffected. Here, the term *local* is used with respect to the \mathcal{A} and \mathcal{B}

partitioning. Specifically, via the local unitary transformation

$$V_{ij} \equiv \begin{cases} \frac{1}{2}(\mathbb{1}_2 + i\sigma_z) \otimes (\mathbb{1}_2 + i\sigma_z) & : i = 1, j = 2, \\ \frac{1}{2}(\sigma_x + \sigma_z) \otimes (\sigma_x + \sigma_z) & : i = 2, j = 3, \\ \frac{1}{2}(\mathbb{1}_2 + i\sigma_z) \otimes (\mathbb{1}_2 - i\sigma_z) & : i = 3, j = 4, \end{cases} \quad (\text{A.4})$$

one can swap Π_i and Π_j while leaving all the other Bell projectors unaffected. In terms of the corresponding 4×4 matrix representation, the effect of such local unitaries on μ amounts to permutation of the rows and/or columns of M . For brevity, in what follows, we will say that two matrices M and M' are local-unitarily equivalent if we can obtain M by simply permuting the rows and/or columns of M' and *vice versa*. A direct consequence of this observation is that if r represents a separable state, so is any other r' that is obtained from r by independently permuting any of its rows and/or columns.

Before we state the main result of this section, let us introduce one more definition.

Definition 18. Let $\mathcal{P}_s \subset \mathcal{R}$ be the convex hull of the states

$$D_0 \equiv \frac{1}{4} \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \quad G_0 \equiv \frac{1}{4} \begin{pmatrix} 1 & 1 & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad (\text{A.5})$$

and the states that are local-unitarily equivalent to these two.

Simple calculations show that with respect to the \mathcal{A} and \mathcal{B} partitioning, D_0, G_0 are separable. In particular, when written in the product basis of $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$, it can be shown that D_0 admits the following convex decomposition in terms of separable states:

$$\begin{aligned} & \frac{1}{8} \left[\left(|0\rangle|0\rangle \otimes |0\rangle|0\rangle + |1\rangle|1\rangle \otimes |1\rangle|1\rangle \right) \left(\langle 0|\langle 0| \otimes \langle 0|\langle 0| + \langle 1|\langle 1| \otimes \langle 1|\langle 1| \right) \right. \\ & + \left(|0\rangle|0\rangle \otimes |1\rangle|1\rangle + |1\rangle|1\rangle \otimes |0\rangle|0\rangle \right) \left(\langle 0|\langle 0| \otimes \langle 1|\langle 1| + \langle 1|\langle 1| \otimes \langle 0|\langle 0| \right) \\ & + \left(|0\rangle|1\rangle \otimes |0\rangle|1\rangle + |1\rangle|0\rangle \otimes |1\rangle|0\rangle \right) \left(\langle 0|\langle 1| \otimes \langle 0|\langle 1| + \langle 1|\langle 0| \otimes \langle 1|\langle 0| \right) \\ & \left. + \left(|0\rangle|1\rangle \otimes |1\rangle|0\rangle + |1\rangle|0\rangle \otimes |0\rangle|1\rangle \right) \left(\langle 0|\langle 1| \otimes \langle 1|\langle 0| + \langle 1|\langle 0| \otimes \langle 0|\langle 1| \right) \right]. \end{aligned} \quad (\text{A.6})$$

Likewise, it can be shown that G_0 admits the following convex decomposition in terms of product states:

$$\begin{aligned} & \frac{1}{4} \left(|0\rangle|0\rangle \langle 0|\langle 0| \otimes |0\rangle|0\rangle \langle 0|\langle 0| + |0\rangle|1\rangle \langle 0|\langle 1| \otimes |0\rangle|1\rangle \langle 0|\langle 1| \right. \\ & \left. + |1\rangle|0\rangle \langle 1|\langle 0| \otimes |1\rangle|0\rangle \langle 1|\langle 0| + |1\rangle|1\rangle \langle 1|\langle 1| \otimes |1\rangle|1\rangle \langle 1|\langle 1| \right). \end{aligned} \quad (\text{A.7})$$

Hence, \mathcal{P}_s is a separable subset of \mathcal{R} . The main result of this section consists of showing the converse, and hence the following theorem.

Theorem 19. \mathcal{P}_s is the set of states in \mathcal{R} that are separable with respect to the \mathcal{A} and \mathcal{B} partitioning.

Now, we note that \mathcal{P}_s is a convex polytope. Its boundary is therefore described by a finite number of facets [106]. Hence, to prove the above theorem, it suffices to show that all these facets correspond to valid entanglement witnesses. Denoting the set of facets by $\mathcal{W} = \{W_i\}$. Then, using the software PORTA,² the *nontrivial* facets were found to be equivalent under local unitaries to one of the following:

$$\begin{aligned} W_1 &\equiv \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}, \quad W_2 \equiv \begin{pmatrix} 1 & 1 & \cdot & -1 \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot \end{pmatrix}, \\ W_3 &\equiv \begin{pmatrix} 3 & 3 & 1 & -1 \\ 3 & -1 & 1 & 3 \\ 1 & 1 & 3 & 1 \\ -1 & -1 & 1 & -1 \end{pmatrix}, \quad W_4 \equiv \begin{pmatrix} 3 & 3 & 1 & -1 \\ 3 & -1 & 1 & 3 \\ 3 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}. \end{aligned} \quad (\text{A.8})$$

Apart from these, there is also a facet W_0 whose only nonzero entry is $[W_0]_{11} = 1$. W_0 and the operators local-unitarily equivalent to it give rise to positive definite matrices [c.f. Eq. (A.9)], and thus correspond to trivial entanglement witnesses. On the other hand, it is also not difficult to verify that W_1 (and operators equivalent under local unitaries) are decomposable and therefore demand that ρ_s remains positive semidefinite after partial transposition. These are all the entanglement witnesses that arise from the positive partial transposition requirement [161, 162] for separable states.

To complete the proof of Theorem 19, it remains to show that W_2, W_3, W_4 give rise to Hermitian matrices

$$Z_{w,k} = \sum_{i=1}^4 \sum_{j=1}^4 [W_k]_{i,j} (\Pi_i \otimes \Pi_j) \quad (\text{A.9})$$

that are valid entanglement witnesses, i.e., $\text{tr}(\rho_s Z_{w,k}) \geq 0$ for any separable $\rho_s \in \mathcal{R}$. It turns out that this can be proved with the help of the following lemma from Ref. [171, 172].

Lemma 20. *For a given Hermitian matrix Z_w acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, with $\dim(\mathcal{H}_A) = d_A$ and $\dim(\mathcal{H}_B) = d_B$, if there exists $m, n \in \mathbb{Z}^+$, positive semidefinite \mathcal{Z} acting on $\mathcal{H}_A^{\otimes m} \otimes \mathcal{H}_B^{\otimes n}$ and a subset s of the $m + n$ tensor factors such that*

$$\Pi_A \otimes \Pi_B \left(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1} \right) \Pi_A \otimes \Pi_B = \Pi_A \otimes \Pi_B \left(\mathcal{Z}^{T_s} \right) \Pi_A \otimes \Pi_B, \quad (\text{A.10})$$

where Π_A is the projector onto the symmetric subspace of $\mathcal{H}_A^{\otimes m}$ (likewise for Π_B) and $(\cdot)^{T_s}$ refers to partial transposition with respect to the subsystem s , then Z_w is a valid entanglement witness across \mathcal{H}_A and \mathcal{H}_B , i.e., $\text{tr}(\rho_{sep} Z_w) \geq 0$ for any state ρ_{sep} that is separable with respect to the \mathcal{A} and \mathcal{B} partitioning.

²This software package, which stands for POlyhedron Representation Transformation Algorithm, is available at <http://www.zib.de/Optimization/Software/Porta/>

Proof. Denote by $\mathcal{A}^{[k]}$ the subsystem associated with the k -th copy of $\mathcal{H}_{\mathcal{A}}$ in $\mathcal{H}_{\mathcal{A}}^{\otimes m}$; likewise for $\mathcal{B}^{[l]}$. To prove the above lemma, let $|\alpha\rangle \in \mathcal{H}_{\mathcal{A}}$ and $|\beta\rangle \in \mathcal{H}_{\mathcal{B}}$ be (unit) vectors, and for definiteness, let $s = \mathcal{B}^{[n]}$ then it follows that

$$\begin{aligned}
& \langle \alpha | \langle \beta | Z_w | \alpha \rangle | \beta \rangle \\
&= \langle \alpha |^{\otimes m} \langle \beta |^{\otimes n} (\mathbf{1}_{d_{\mathcal{A}}}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_{\mathcal{B}}}^{\otimes n-1}) | \alpha \rangle^{\otimes m} | \beta \rangle^{\otimes n} \\
&= \langle \alpha |^{\otimes m} \langle \beta |^{\otimes n} [\Pi_{\mathcal{A}} \otimes \Pi_{\mathcal{B}} (\mathcal{Z}^{T_s}) \Pi_{\mathcal{A}} \otimes \Pi_{\mathcal{B}}] | \alpha \rangle^{\otimes m} | \beta \rangle^{\otimes n} \\
&= \langle \alpha |^{\otimes m} \langle \beta |^{\otimes n} (\mathcal{Z}^{T_{\mathcal{B}^{[n]}}}) | \alpha \rangle^{\otimes m} | \beta \rangle^{\otimes n} \\
&= (\langle \alpha |^{\otimes m} \langle \beta |^{\otimes n-1} \otimes \langle \beta^* |) \mathcal{Z} (| \alpha \rangle^{\otimes m} | \beta \rangle^{\otimes n-1} \otimes | \beta^* \rangle) \\
&\geq 0,
\end{aligned}$$

where $|\beta^*\rangle$ is the complex conjugate of $|\beta\rangle$. We have made use of the identity $\Pi_{\mathcal{A}}|\alpha\rangle^{\otimes m} = |\alpha\rangle^{\otimes m}$ (likewise for $\Pi_{\mathcal{B}}$) in the second and third equality, Eq. (A.10) in the second equality, and the positive semidefiniteness of \mathcal{Z} . To cater for general s , we just have to modify the second to last line of the above computation accordingly (i.e., to perform complex conjugation on all the states in the set s) and the proof will proceed as before. \square

More generally, let us remark that instead of having one \mathcal{Z} on the *rhs* of Eq. (A.10), one can also have a sum of different \mathcal{Z} 's, with each of them partial transposed with respect to different subsystems s . Clearly, if the given Z_w admits such a decomposition, it is also an entanglement witness [171, 172]. For our purposes these more complicated decompositions do not offer any advantage over the simple decomposition given in Eq. (A.10).

By solving some appropriate SDPs (Appendix C.3.3), we have found that when $m = 3$, $n = 2$ and $s = \mathcal{B}^{[2]}$, there exist some $\mathcal{Z}_k \geq 0$, such that Eq. (A.10) holds true for each $k \in \{1, 2, 3, 4\}$. Due to space limitations, the analytic expression for these \mathcal{Z}_k 's will not be reproduced here but are made available online at Ref. [220]. For W_2 , the fact that the corresponding $Z_{w,2}$ is a witness can even be verified by considering $m = 2$, $n = 1$ and $s = \mathcal{A}^{[1]}$. In this case, $d_{\mathcal{A}} = d_{\mathcal{B}} = 4$. If we label the local basis vectors by $\{|i\rangle\}_{i=0}^3$, the corresponding \mathcal{Z} reads

$$\begin{aligned}
\mathcal{Z}_2 &= \frac{1}{2} \sum_{i=1}^4 |z_i\rangle\langle z_i|, \\
|z_1\rangle &= |01, 0\rangle - |02, 3\rangle + |11, 1\rangle + |13, 3\rangle + |22, 1\rangle + |23, 0\rangle, \\
|z_2\rangle &= |10, 3\rangle + |11, 2\rangle + |20, 0\rangle + |22, 2\rangle - |31, 0\rangle + |32, 3\rangle, \\
|z_3\rangle &= |00, 0\rangle + |02, 2\rangle + |10, 1\rangle - |13, 2\rangle + |32, 1\rangle + |33, 0\rangle, \\
|z_4\rangle &= |00, 3\rangle + |01, 2\rangle - |20, 1\rangle + |23, 2\rangle + |31, 1\rangle + |33, 3\rangle,
\end{aligned}$$

where we have separated \mathcal{A} 's degree of freedom from \mathcal{B} 's ones by comma.³ This completes the proof for Theorem 19.

An immediate corollary of the above characterization is that we now know exactly the set of Bell-diagonal preserving transformations that can be performed locally on a Bell-diagonal

³Note that to verify \mathcal{Z}_2 against Eq. (A.10), one should also rewrite $Z_{w,2}$ obtained in Eq. (A.9) in the appropriate tensor-product basis such that $Z_{w,2}$ acts on $\mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{A}''} \otimes \mathcal{H}_{\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{B}''}$.

state. In what follows, we will make use of the Choi-Jamiołkowski isomorphism [221, 222, 223], i.e., the one-to-one correspondence between CPM and quantum state, to make these SLOCC transformations explicit.

A.2 Separable Maps and SLOCC

Now, let us recall some well-established facts about CPM. To begin with, a separable CPM, denoted by \mathcal{E}_s takes the following form [224, 225]

$$\mathcal{E}_s : \rho \rightarrow \sum_{i=1}^n (A_i \otimes B_i) \rho (A_i^\dagger \otimes B_i^\dagger), \quad (\text{A.11})$$

where ρ acts on $\mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{in}}}$, A_i acts on $\mathcal{H}_{\mathcal{A}_{\text{in}}}$, B_i acts on $\mathcal{H}_{\mathcal{B}_{\text{in}}}$.⁴

If, moreover,

$$\sum_i (A_i \otimes B_i)^\dagger (A_i \otimes B_i) = \mathbb{1}, \quad (\text{A.12})$$

the map is trace-preserving, i.e., if ρ is normalized, so is the output of the map $\mathcal{E}_s(\rho)$. Equivalently, the trace-preserving condition demands that the transformation from ρ to $\mathcal{E}_s(\rho)$ can always be achieved with certainty. It is well-known that all LOCC transformations are of the form Eq. (A.11) but the converse is not true [226].

However, if we allow the map $\rho \rightarrow \mathcal{E}_s(\rho)$ to fail with some probability $p < 1$, the transformation from ρ to $\mathcal{E}_s(\rho)$ can always be implemented probabilistically via LOCC. In other words, if we do not impose Eq. (A.12), then Eq. (A.11) represents, up to some normalization constant, the most general LOCC possible on a bipartite quantum system. These are the SLOCC transformations [213].

To see that Eq. (A.11) can always be realized with some non-zero probability of success, we first note that each of the terms in the decomposition can always be implemented with some probability of success. For instance, if they wish to implement the k^{th} term in Eq. (A.11), i.e., $(A_k \otimes B_k) \rho (A_k^\dagger \otimes B_k^\dagger)$ — which by itself represents uncorrelated local quantum operations on the individual subsystems, they can do that by just by applying some local unitary transformation and/or measurement on their local subsystem. With the help of classical communication, they can then postselect on the desired outcomes to achieve the transformation $(A_k \otimes B_k) \rho (A_k^\dagger \otimes B_k^\dagger)$.

With that in mind, it is then easy to see that implementation of the separable map can be carried out by probabilistically selecting the term to implement in the separable map given by Eq. (A.11). Party \mathcal{A} can first toss a coin to decide on the term in the decomposition [c.f. Eq. (A.11)] that she would like to implement for that run of the experiment and communicate this outcome to Bob. They then both perform appropriate local operations and postselection to achieve the desired transformation with some probability. Clearly, since each term in Eq. (A.11) can be implemented with some non-zero probability of success, so can the separable map given by Eq. (A.11).

⁴Following Kraus' work on CPM [211, 212], this specific form of the CPM is also known as a Kraus decomposition of the CPM, with each $A_i \otimes B_i$ in the sum conventionally called the Kraus operator associated with the CPM.

Now, to make a connection between the set of SLOCC transformations and the set of states that we have characterized in Sec. A.1, let us also recall the Choi-Jamiołkowski isomorphism [221, 222, 223] between CPM and quantum states: for every (not necessarily separable) CPM $\mathcal{E} : \mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{in}}} \rightarrow \mathcal{H}_{\mathcal{A}_{\text{out}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{out}}}$ there is a unique — again, up to some positive constant α — quantum state $\rho_{\mathcal{E}}$ corresponding to \mathcal{E} :

$$\rho_{\mathcal{E}} = \alpha \mathcal{E}_{\text{in}} \otimes \mathcal{I}_{\text{out}} (|\Phi^+\rangle_{\mathcal{A}\mathcal{A}} \langle \Phi^+| \otimes |\Phi^+\rangle_{\mathcal{B}\mathcal{B}} \langle \Phi^+|), \quad (\text{A.13})$$

where $|\Phi^+\rangle_{\mathcal{A}} \equiv \sum_{i=1}^{d_{\mathcal{A}_{\text{in}}}} |i\rangle_{\text{in}} \otimes |i\rangle_{\text{out}}$ is the unnormalized maximally entangled state of dimension $d_{\mathcal{A}_{\text{in}}}$ (likewise for $|\Phi^+\rangle_{\mathcal{B}}$). In Eq. (A.13), it is understood that \mathcal{E}_{in} only acts on the “in” space of $|\Phi^+\rangle_{\mathcal{A}}$ and $|\Phi^+\rangle_{\mathcal{B}}$. Clearly, the state $\rho_{\mathcal{E}}$ acts on a Hilbert space of dimension $d_{\mathcal{A}_{\text{in}}} \times d_{\mathcal{A}_{\text{out}}} \times d_{\mathcal{B}_{\text{in}}} \times d_{\mathcal{B}_{\text{out}}}$, where $d_{\mathcal{A}_{\text{out}}} \times d_{\mathcal{B}_{\text{out}}}$ is the dimension of $\mathcal{H}_{\mathcal{A}_{\text{out}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{out}}}$.

Conversely, given a state $\rho_{\mathcal{E}}$ acting on $\mathcal{H}_{\mathcal{A}_{\text{out}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{out}}} \otimes \mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{in}}}$, the corresponding action of the CPM \mathcal{E} on some ρ acting on $\mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{in}}}$ reads:

$$\mathcal{E}(\rho) = \frac{1}{\alpha} \text{tr}_{\mathcal{A}_{\text{in}}\mathcal{B}_{\text{in}}} [\rho_{\mathcal{E}} (\mathbb{1}_{\mathcal{A}_{\text{out}}\mathcal{B}_{\text{out}}} \otimes \rho^{\text{T}})], \quad (\text{A.14})$$

where ρ^{T} denotes transposition of ρ in some local bases of $\mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{in}}}$. For a trace-preserving CPM, it then follows that we must have $\text{tr}_{\mathcal{A}_{\text{out}}\mathcal{B}_{\text{out}}}(\rho_{\mathcal{E}}) = \alpha \mathbb{1}_{\mathcal{A}_{\text{in}}\mathcal{B}_{\text{in}}}$. A point that should be emphasized now is that \mathcal{E} is a separable map, Eq. (A.11), if and only if the corresponding $\rho_{\mathcal{E}}$ given by Eq. (A.13) is separable across $\mathcal{H}_{\mathcal{A}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{A}_{\text{out}}}$ and $\mathcal{H}_{\mathcal{B}_{\text{in}}} \otimes \mathcal{H}_{\mathcal{B}_{\text{out}}}$ [227]. Moreover, at the risk of repeating ourselves, the map $\rho \rightarrow \mathcal{E}(\rho)$ derived from a separable $\rho_{\mathcal{E}}$ can always be implemented locally, although it may only succeed with some (nonzero) probability. Hence, if we are only interested in transformations that can be performed locally, and not the probability of success in mapping $\rho \rightarrow \mathcal{E}(\rho)$, the normalization constant α as well as the normalization of $\rho_{\mathcal{E}}$ becomes irrelevant. This is the convention that we will adopt for the rest of this Appendix.

A.3 Bell-diagonal Preserving SLOCC Transformations

We shall now apply the isomorphism to the class of states \mathcal{R} that we have characterized in Sec. A.1. In particular, if we identify \mathcal{A}_{in} , \mathcal{A}_{out} , \mathcal{B}_{in} and \mathcal{B}_{out} with, respectively, \mathcal{A}'' , \mathcal{A}' , \mathcal{B}'' and \mathcal{B}' , it follows from Eq. (A.2) and Eq. (A.14) that for any two-qubit state ρ_{in} , the action of the CPM derived from $\rho \in \mathcal{R}$ reads:

$$\mathcal{E} : \rho_{\text{in}} \rightarrow \rho_{\text{out}} \propto \sum_{i,j} [r]_{i,j} \text{tr}(\rho_{\text{in}}^{\text{T}} \Pi_j) \Pi_i. \quad (\text{A.15})$$

Hence, under the action of \mathcal{E} , any ρ_{in} is transformed to another two-qubit state that is diagonal in the Bell basis, i.e., a Bell-diagonal state. In particular, for a Bell-diagonal ρ_{in} , i.e.,

$$\begin{aligned} \rho_{\text{in}} &= \sum_k [\beta]_k \Pi_k, \\ [\beta]_k &\geq 0, \quad \sum_k [\beta]_k = 1, \end{aligned} \quad (\text{A.16})$$

the map outputs another Bell-diagonal state

$$\rho_{\text{out}} = \mathcal{E}(\rho_{\text{in}}) \propto \sum_{i,j} [\beta]_j [r]_{i,j} \Pi_i. \quad (\text{A.17})$$

It is worth noting that for a general $\rho_{\mathcal{E}} \in \mathcal{R}$, $\text{tr}_{\mathcal{A}'\mathcal{B}'} \rho_{\mathcal{E}}$ is not proportional to the identity matrix, therefore some of the CPMs derived from $\rho \in \mathcal{R}$ are intrinsically non-trace-preserving.⁵

By considering the convex cone⁶ of separable states \mathcal{P}_s that we have characterized in Sec. A.1, we therefore obtain the entire set of Bell-diagonal preserving SLOCC transformations. Among them, we note that the extremal maps, i.e., those derived from Eq. (A.5), admit simple physical interpretations and implementations. In particular, the extremal separable map for D_0 , and the maps that are related to it by local unitaries, correspond to permutation of the input Bell projectors Π_i — which can be implemented by performing appropriate local unitary transformations. The other kind of extremal separable map, derived from G_0 , corresponds to making a measurement that determines if the initial state is in a subspace spanned by a given pair of Bell states and if successful discarding the input state and replacing it by an equal but incoherent mixture of two of the Bell states. This operation can be implemented locally since the equally weighted mixture of two Bell states is a separable state and hence both the measurement step and the state preparation step can be implemented locally.

⁵The $\rho_{\mathcal{E}}$ derived from G_0 in Eq. (A.5) is an example of this sort. In fact, in this case, if the input state has no support on Π_1 nor Π_2 , the map always outputs the zero matrix.

⁶Since the mapping from any $\rho \in \mathcal{P}_s$ to a separable CPM via Eq. (A.14) is only defined up to a positive constant, for the subsequent discussion, we might as well consider the cone generated by \mathcal{P}_s .

B

Some Miscellaneous Calculations

B.1 Classical Correlations and Bell's Theorems

B.1.1 Equivalence between the CGLMP and I_{22nn} inequality

In this section, we will provide a proof that the CGLMP inequality for $n_A = n_B = n$ outcomes, Eq. (3.36), is equivalent to the I_{22nn} inequality, Eq. (3.35). For the purpose of this proof, we will rewrite the CGLMP inequality, Eq. (3.36), by shifting the constant “2” to the *lhs* of the inequality, namely,

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_n)} = & \sum_{k=0}^{\lfloor \frac{n}{2}-1 \rfloor} \left(1 - \frac{2k}{n-1}\right) \sum_{o_b=1}^n \left[p_{\mathcal{AB}}^{o_b-k, o_b}(1, 1) - p_{\mathcal{AB}}^{o_b+k+1, o_b}(1, 1) + p_{\mathcal{AB}}^{o_b+k, o_b}(1, 2) - p_{\mathcal{AB}}^{o_b-k-1, o_b}(1, 2) \right. \\ & \left. + p_{\mathcal{AB}}^{o_b+k, o_b}(2, 1) - p_{\mathcal{AB}}^{o_b-k-1, o_b}(2, 1) + p_{\mathcal{AB}}^{o_b-k-1, o_b}(2, 2) - p_{\mathcal{AB}}^{o_b+k, o_b}(2, 2) \right] - 2 \leq 0, \end{aligned} \quad (\text{B.1})$$

where we remind the reader that expression such as $o_b - k$ in the above inequality is understood to be evaluated modulo n . For ease of reference, we will also reproduce the I_{22nn} inequality as follow:

$$\begin{aligned} \mathcal{S}_{\text{LHV}}^{(I_{22nn})} = & \sum_{o_a=1}^{n-1} \sum_{o_b=1}^{n-o_a} p_{\mathcal{AB}}^{o_a o_b}(1, 1) + \sum_{o_a=1}^{n-1} \sum_{o_b=n-o_a}^{n-1} \left[p_{\mathcal{AB}}^{o_a o_b}(1, 2) + p_{\mathcal{AB}}^{o_a o_b}(2, 1) - p_{\mathcal{AB}}^{o_a o_b}(2, 2) \right] \\ & - \sum_{o_a=1}^{n-1} p_{\mathcal{A}}^{o_a}(1) - \sum_{o_b=1}^{n-1} p_{\mathcal{B}}^{o_b}(1) \leq 0. \end{aligned} \quad (3.35)$$

Moreover, we shall make use of the matrix representation of a Bell inequality for probabilities introduced in Eq. (3.16) – Eq. (3.20).

Let us begin by showing the equivalence explicitly for $n = 3$. In this case, the *lhs* of inequality (B.1) can be represented by the following matrix of coefficients¹, c.f. Eq. (3.19) and Eq. (3.20),

$$b^{(I_3)} : \sim \left(\begin{array}{c|c|c} b_{0,0} & \mathbf{b}_{0,1} & \mathbf{b}_{0,2} \\ \hline \mathbf{b}_{1,0} & \mathbf{b}_{1,1} & \mathbf{b}_{1,2} \\ \hline \mathbf{b}_{2,0} & \mathbf{b}_{2,1} & \mathbf{b}_{2,1} \end{array} \right) = \left(\begin{array}{c|cc|cc|c} -2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 1 & \cdot & -1 & 1 & -1 & \cdot \\ \cdot & -1 & 1 & \cdot & \cdot & 1 & -1 \\ \cdot & \cdot & -1 & 1 & -1 & \cdot & 1 \\ \hline \cdot & 1 & -1 & \cdot & -1 & 1 & \cdot \\ \cdot & \cdot & 1 & -1 & \cdot & -1 & 1 \\ \cdot & -1 & \cdot & 1 & 1 & \cdot & -1 \end{array} \right) \quad (\text{B.2a})$$

where we recall that coefficients associated with Alice's (Bob's) local measurement setting are separated by a single horizontal (vertical) line; coefficients associated with marginal probabilities are separated from the others via double horizontal (vertical) lines.

Now, let us make use of the no-signaling condition, Eq. (3.5) to express the joint probabilities associated with Alice's third measurement outcome $p_{\mathcal{AB}}^{3o_b}(s_a, s_b)$ in terms of marginal probabilities $p_{\mathcal{B}}^{o_b}(s_b)$ and the other joint probabilities $p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b)$ for $o_a \neq 3$. For example, doing this for $s_a = s_b = 1$, $o_b = 2$ amounts to subtracting every entry in the second column of the block matrix $\mathbf{b}_{1,1}$ by the entry $[\mathbf{b}_{1,1}]_{3,2}$ and adding this specific entry to the marginal entry $[\mathbf{b}_{0,1}]_2$ directly above it. Repeating this for all combinations of s_a , s_b and o_b gives rise to an equivalent inequality with matrix of coefficients given by

$$b'^{(I_3)} : \sim \left(\begin{array}{c|ccc|ccc} -2 & -1 & -1 & 2 & \cdot & \cdot & \cdot \\ \hline \cdot & 1 & 1 & -2 & 2 & -1 & -1 \\ \cdot & -1 & 2 & -1 & 1 & 1 & -2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 2 & -1 & -1 & -2 & 1 & 1 \\ \cdot & 1 & 1 & -2 & -1 & -1 & 2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right). \quad (\text{B.2b})$$

Next, let us also make use of the no-signaling condition to express the joint probabilities associated with Bob's third measurement outcome $p_{\mathcal{AB}}^{o_a 3}(s_a, s_b)$ in terms of marginal probabilities $p_{\mathcal{A}}^{o_a}(s_a)$ and the other joint probabilities $p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b)$ for $o_b \neq 3$. In particular, for $s_a = 1, s_b = o_b = 2$, this amounts to subtracting every entry in the second row of the block matrix $\mathbf{b}'_{1,2}$ by the entry $[\mathbf{b}'_{1,2}]_{2,3}$ and adding this specific entry to the marginal entry $[\mathbf{b}'_{1,0}]_2$ that is on the same row. Doing this for all combinations of s_a , s_b and o_a gives rise to another

¹Notice that here, we are writing each block matrix in full dimension, i.e., each block is of dimension $n_A \times n_B$.

equivalent inequality with matrix of coefficients given by

$$b''(I_3) \sim \left(\begin{array}{c|ccc|ccc} -2 & -1 & -1 & 2 & \cdot & \cdot & \cdot \\ \hline -3 & 3 & 3 & \cdot & 3 & \cdot & \cdot \\ -3 & \cdot & 3 & \cdot & 3 & 3 & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 3 & \cdot & \cdot & -3 & \cdot & \cdot \\ \cdot & 3 & 3 & \cdot & -3 & -3 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right). \quad (\text{B.2c})$$

Then, by using the normalization condition, c.f. Eq. (3.4),

$$\sum_{o_b=1}^n p_B^{o_b}(s_b) = 1, \quad (\text{B.2d})$$

we can further express $p_B^3(1)$ in terms of $p_B^{o_b}(1)$ for $o_b \neq 3$. In terms of the matrix of coefficients, this amounts to subtracting every entry in $\mathbf{b}_{0,1}''$ by $[\mathbf{b}_{0,1}'']_3$ and adding this specific entry to $b_{0,0}''$. Writing this out explicitly, we get the matrix of coefficients for a fourth equivalent inequality:

$$b'''(I_3) \sim \left(\begin{array}{c|ccc|ccc} \cdot & -3 & -3 & \cdot & \cdot & \cdot & \cdot \\ \hline -3 & 3 & 3 & \cdot & 3 & \cdot & \cdot \\ -3 & \cdot & 3 & \cdot & 3 & 3 & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 3 & \cdot & \cdot & -3 & \cdot & \cdot \\ \cdot & 3 & 3 & \cdot & -3 & -3 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right). \quad (\text{B.2e})$$

What remains to be done now is to swap all of Bob's first and second measurement outcomes, which gives a fifth equivalent inequality with matrix of coefficients:

$$b''''(I_3) \sim 3 \left(\begin{array}{c|ccc|ccc} \cdot & -1 & -1 & \cdot & \cdot & \cdot & \cdot \\ \hline -1 & 1 & 1 & \cdot & \cdot & 1 & \cdot \\ -1 & 1 & \cdot & \cdot & 1 & 1 & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & 1 & \cdot & \cdot & -1 & \cdot \\ \cdot & 1 & 1 & \cdot & -1 & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right), \quad (\text{B.2f})$$

which can be seen to be equivalent to the I_{2233} inequality, Eq. (3.32).

In exactly the same manner, we see that for $n = 4$, we have

$$b^{(I_4)} : \sim \left(\begin{array}{c|cccc|cccc} -2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 1 & 1/3 & -1/3 & -1 & 1 & -1 & -1/3 & 1/3 \\ \cdot & -1 & 1 & 1/3 & -1/3 & 1/3 & 1 & -1 & -1/3 \\ \cdot & -1/3 & -1 & 1 & 1/3 & -1/3 & 1/3 & 1 & -1 \\ \cdot & 1/3 & -1/3 & -1 & 1 & -1 & -1/3 & 1/3 & 1 \\ \hline \cdot & 1 & -1 & -1/3 & 1/3 & -1 & 1 & 1/3 & -1/3 \\ \cdot & 1/3 & 1 & -1 & -1/3 & -1/3 & -1 & 1 & 1/3 \\ \cdot & -1/3 & 1/3 & 1 & -1 & 1/3 & -1/3 & -1 & 1 \\ \cdot & -1 & -1/3 & 1/3 & 1 & 1 & 1/3 & -1/3 & -1 \end{array} \right),$$

then, by zeroing the coefficients associated with $p_{AB}^{4o_b}(s_a, s_b)$ and $p_{AB}^{o_a4}(s_a, s_b)$ via the no-signaling condition, Eq. (3.5), we get

$$\Rightarrow b''(I_4) : \sim \begin{pmatrix} -2 & -2/3 & -2/3 & -2/3 & 2 & . & . & . & . \\ \hline . & 2/3 & 2/3 & 2/3 & -2 & 2 & -2/3 & -2/3 & -2/3 \\ . & -4/3 & 4/3 & 4/3 & -4/3 & 4/3 & 4/3 & -4/3 & -4/3 \\ . & -2/3 & -2/3 & 2 & -2/3 & 2/3 & 2/3 & 2/3 & -2 \\ . & . & . & . & . & . & . & . & . \\ \hline . & 2 & -2/3 & -2/3 & -2/3 & -2 & 2/3 & 2/3 & 2/3 \\ . & 4/3 & 4/3 & -4/3 & -4/3 & -4/3 & -4/3 & 4/3 & 4/3 \\ . & 2/3 & 2/3 & 2/3 & -2 & -2/3 & -2/3 & -2/3 & 2 \\ . & . & . & . & . & . & . & . & . \end{pmatrix}$$

$$\Rightarrow b''(I_4) : \sim \begin{pmatrix} -2 & -2/3 & -2/3 & -2/3 & 2 & . & . & . & . \\ \hline -8/3 & 8/3 & 8/3 & 8/3 & . & 8/3 & . & . & . \\ -8/3 & . & 8/3 & 8/3 & . & 8/3 & 8/3 & . & . \\ -8/3 & . & . & 8/3 & . & 8/3 & 8/3 & 8/3 & . \\ . & . & . & . & . & . & . & . & . \\ \hline . & 8/3 & . & . & . & -8/3 & . & . & . \\ . & 8/3 & 8/3 & . & . & -8/3 & -8/3 & . & . \\ . & 8/3 & 8/3 & 8/3 & . & -8/3 & -8/3 & -8/3 & . \\ . & . & . & . & . & . & . & . & . \end{pmatrix}.$$

Finally, by zeroing $p_B^4(1)$ using Eq. (B.2d) and swapping Bob’s measurement outcomes, $o_b \leftrightarrow n - o_b$, we end up with an equivalent inequality with matrix of coefficients:

$$b''''(I_4) : \sim_{\mathbb{Z}}^8 \begin{pmatrix} . & -1 & -1 & -1 & . \\ \hline -1 & 1 & 1 & 1 & . \\ -1 & 1 & 1 & . & . \\ -1 & 1 & . & . & . \\ . & . & . & . & . \\ \hline . & . & . & 1 & . \\ . & . & 1 & 1 & . \\ . & 1 & 1 & 1 & . \\ . & . & . & . & . \end{pmatrix},$$

I_4 is thus equivalent to the I_{2244} inequality, c.f. Eq. (3.34).

More generally, we can prove that I_n is equivalent to I_{22nn} by generalizing the above procedures. Firstly, we note that the matrix of coefficients $b^{(I_n)}$ for the n -outcome CGLMP inequality, Eq. (B.1), is made up from blocks of circulant matrices \mathbf{b}_{s_a, s_b} , with entries given by

$$[\mathbf{b}_{1,1}]_{o_a, o_b} = \begin{cases} 1 + \frac{2(o_a - o_b)}{n-1} & : o_b \geq o_a \\ -1 + \frac{2(o_a - o_b - 1)}{n-1} & : o_b < o_a \end{cases}, \quad (\text{B.3a})$$

$$[\mathbf{b}_{1,2}]_{o_a, o_b} = [\mathbf{b}_{2,1}]_{o_a, o_b} = -[\mathbf{b}_{2,2}]_{o_a, o_b} = \begin{cases} -1 + \frac{2(o_b - o_a - 1)}{n-1} & : o_b > o_a \\ 1 + \frac{2(o_b - o_a)}{n-1} & : o_b \leq o_a \end{cases}, \quad (\text{B.3b})$$

and marginal blocks $\mathbf{b}_{s_a, 0}$, \mathbf{b}_{0, s_b} , $b_{0,0}$:

$$\mathbf{b}_{1,0} = \mathbf{b}_{2,0} = \mathbf{0}_n, \quad \mathbf{b}_{0,1} = \mathbf{b}_{0,2} = \mathbf{0}_n^T, \quad b_{0,0} = -2. \quad (\text{B.3c})$$

where $\mathbf{0}_n$ is an $n \times 1$ null vector.

As in Eq. (B.2b), we will now make use of the no-signaling condition, Eq. (3.5), to zero the coefficients associated with the joint probabilities $p_{AB}^{n o_b}(s_a, s_b)$. This gives rise to an equivalent inequality whose matrix of coefficients $b'^{(I_n)}$ is related to the original one, $b^{(I_n)}$ by

$$\begin{aligned} \mathbf{b}'_{0,0} &= \mathbf{b}_{0,0}, \quad \mathbf{b}'_{s_a,0} = \mathbf{b}_{s_a,0}, \quad s_a = 1, 2, \\ [\mathbf{b}'_{0,s_b}]_{o_b} &= [\mathbf{b}_{0,s_b}]_{o_b} + \sum_{s_a=1}^2 [\mathbf{b}_{s_a,s_b}]_{n,o_b}, \quad s_b = 1, 2, \quad o_b = 1, 2, \dots, n, \\ [\mathbf{b}'_{s_a,s_b}]_{o_a, o_b} &= [\mathbf{b}_{s_a,s_b}]_{o_a, o_b} - [\mathbf{b}_{s_a,s_b}]_{n,o_b}, \quad s_a, s_b = 1, 2, \quad o_a, o_b = 1, 2, \dots, n. \end{aligned}$$

Next, we will again make use of the no-signaling condition, but to instead zero the coefficients associated with the joint probabilities $p_{AB}^{o_a n}(s_a, s_b)$ [c.f. Eq. (B.2c)]. This gives rise to another equivalent inequality whose matrix of coefficients $b''^{(I_n)}$ is related to the existing one, $b'^{(I_n)}$ by

$$\begin{aligned} \mathbf{b}''_{0,0} &= \mathbf{b}'_{0,0}, \quad \mathbf{b}''_{0,s_b} = \mathbf{b}'_{0,s_b}, \quad s_b = 1, 2, \\ [\mathbf{b}''_{s_a,0}]_{o_a} &= [\mathbf{b}'_{s_a,0}]_{o_a} + \sum_{s_b=1}^2 [\mathbf{b}'_{s_a,s_b}]_{o_a,n}, \quad s_a = 1, 2, \quad o_a = 1, 2, \dots, n, \\ [\mathbf{b}''_{s_a,s_b}]_{o_a, o_b} &= [\mathbf{b}'_{s_a,s_b}]_{o_a, o_b} - [\mathbf{b}'_{s_a,s_b}]_{o_a,n}, \quad s_a, s_b = 1, 2, \quad o_a, o_b = 1, 2, \dots, n. \end{aligned}$$

Now, we will make use of the normalization of marginal probabilities, Eq. (B.2d), to zero the coefficients associated with the marginal probabilities $p_A^n(s_a)$. This gives rise to another equivalent inequality whose matrix of coefficients $b'''^{(I_n)}$ is related to the existing one, $b''^{(I_n)}$ by

$$\begin{aligned} b'''_{0,0} &= b''_{0,0} + \sum_{s_b=1}^2 [\mathbf{b}''_{0,s_b}]_n, \\ [\mathbf{b}'''_{0,s_b}]_{o_b} &= [\mathbf{b}''_{0,s_b}]_{o_b} - [\mathbf{b}''_{0,s_b}]_n, \quad s_b = 1, 2, \quad o_b = 1, 2, \dots, n, \\ \mathbf{b}'''_{s_a,0} &= \mathbf{b}''_{s_a,0}, \quad s_a = 1, 2, \\ \mathbf{b}'''_{s_a,s_b} &= \mathbf{b}''_{s_a,s_b}, \quad s_a, s_b = 1, 2. \end{aligned}$$

More explicitly, it is easy to check that the matrix of coefficients $b'''^{(I_n)}$ reads

$$[\mathbf{b}_{1,1}''']_{o_a, o_b} = \begin{cases} \frac{2n}{n-1} & : o_b \geq o_a, o_b \neq n \\ 0 & : o_b \geq o_a, o_b = n \\ 0 & : o_b < o_a, \end{cases} \quad (\text{B.4a})$$

$$[\mathbf{b}_{1,2}''']_{o_a, o_b} = [\mathbf{b}_{2,1}''']_{o_a, o_b} = -[\mathbf{b}_{2,2}''']_{o_a, o_b} = \begin{cases} 0 & : o_b > o_a \\ \frac{2n}{n-1} & : o_b \leq o_a, o_a \neq n \\ 0 & : o_b \leq o_a, o_a = n, \end{cases} \quad (\text{B.4b})$$

$$[\mathbf{b}_{0,1}''']_{o_b} = \begin{cases} -\frac{2n}{n-1} & : o_b < n \\ 0 & : o_b = n \end{cases} \quad (\text{B.4c})$$

$$[\mathbf{b}_{1,0}''']_{o_a} = \begin{cases} -\frac{2n}{n-1} & : o_a < n \\ 0 & : o_a = n \end{cases} \quad (\text{B.4d})$$

$$\mathbf{b}_{2,0}''' = \mathbf{0}_n, \quad \mathbf{b}_{0,2}''' = \mathbf{0}_n^\top, \quad \mathbf{b}_{0,0}''' = 0. \quad (\text{B.4e})$$

Finally, by swapping Bob's measurement outcomes, $o_b \leftrightarrow n - o_b$ for all o_b , s_a and s_b , it is readily seen that the matrix of coefficients for this fifth equivalent inequality is related to that of I_{22nn} by

$$b''''^{(I_n)} = \frac{2n}{n-1} b^{(I_{22nn})}, \quad (\text{B.5})$$

and hence the CGLMP inequality with n outcomes, Eq. (B.1), is equivalent to the I_{22nn} inequality. Clearly, since the two inequalities, Eq. (3.36) and Eq. (B.1), are identical up to simple algebraic manipulations, the I_{22nn} inequality must also be equivalent to the CGLMP inequality written in the form of Eq. (3.36). In particular, if we denote by \mathcal{B}_{I_n} and $\mathcal{B}_{I_{22nn}}$, respectively, the Bell operator derived from Eq. (3.36) and Eq. (3.35). Then, for any quantum state ρ , the expectation values of these Bell operators with respect to ρ are related by

$$\text{tr}(\rho \mathcal{B}_{I_n}) = \frac{2n}{n-1} \text{tr}(\rho \mathcal{B}_{I_{22nn}}) + 2. \quad (6.8)$$

B.2 Quantum Correlations and Locally Causal Quantum States

B.2.1 Convexity of Non-Bell-Inequality-Violating States

Here, we will prove that the set of quantum states not violating a given Bell inequality in the sense of Definition 7 is convex. Let us denote by \mathcal{B}_k the Bell operator associated with a Bell inequality $I_k : \mathcal{S}_{\text{LHV}}^{(k)} \leq \beta_{\text{LHV}}^{(k)}$ and $\mathcal{NV}^{(k)}$ the set of quantum states not violating I_k via a standard Bell experiment. A quantum state $\rho \in \mathcal{NV}^{(k)}$ if

$$\mathcal{S}_{\text{QM}}^{(k)}(\rho) \leq \beta_{\text{LHV}}^{(k)}. \quad (\text{B.6})$$

Clearly, this implies that for any local observables constituting the Bell operator \mathcal{B}_k , we must have

$$\text{tr}(\rho \mathcal{B}_k) \leq \beta_{\text{LHV}}^{(k)}. \quad (\text{B.7})$$

Now, let two quantum states $\rho_1, \rho_2 \in \mathcal{NV}^{(k)}$. Then, for any convex combination of them, i.e.,

$$\rho' \equiv p \rho_1 + (1 - p) \rho_2, \quad 0 \leq p \leq 1, \quad (\text{B.8})$$

we see that

$$\begin{aligned} \text{tr}(\rho' \mathcal{B}_k) &= p \text{tr}(\rho_1 \mathcal{B}_k) + (1 - p) \text{tr}(\rho_2 \mathcal{B}_k), \\ &\leq p \beta_{\text{LHV}}^{(k)} + (1 - p) \beta_{\text{LHV}}^{(k)}, \\ &= \beta_{\text{LHV}}^{(k)}. \end{aligned}$$

Since this is true for any local observables constituting \mathcal{B}_k , we must have

$$\mathcal{S}_{\text{QM}}^{(k)}(\rho') \leq \beta_{\text{LHV}}^{(k)}. \quad (\text{B.9})$$

That is, ρ' is also a member of $\mathcal{NV}^{(k)}$, and hence $\mathcal{NV}^{(k)}$ is a convex set. Moreover, since this is true for an arbitrary Bell inequality, it follows that the set of NBIV quantum states, \mathcal{NV} is also convex.

B.3 Bounds on Quantum Correlations in Standard Bell Experiments

B.3.1 Bell-CH Inequality and Full Rank Projector

In this section, we will prove that the Bell-CH inequality with only two possible outcomes cannot be violated if any of the POVM elements involved is a full rank projector.² For that matter, it suffices to show that in any of these scenarios, the resulting Bell operator \mathcal{B}_{CH} is strictly negative semidefinite (NSD), since the trace of a positive semidefinite (PSD) matrix ρ against an NSD matrix cannot be positive.

Now, let us recall that for a two-outcome Bell experiment, the Bell-CH operator can be written as

$$\mathcal{B}_{\text{CH}} = A_1^+ \otimes (B_2^+ - B_1^-) - A_2^- \otimes B_1^+ - A_2^+ \otimes B_2^+. \quad (5.39)$$

Then, by utilizing the normalization of POVM elements, Eq. (4.1a), we see that when

$$1. \ B_1^+ = \mathbf{0}_{d_{\mathcal{B}} \times d_{\mathcal{B}}}, \ B_1^- = \mathbf{1}_{d_{\mathcal{B}}},$$

$$\mathcal{B}_{\text{CH}} = -A_1^+ \otimes B_2^- - A_2^+ \otimes B_2^+ \leq \mathbf{0}_{d_{\mathcal{A}} d_{\mathcal{B}} \times d_{\mathcal{A}} d_{\mathcal{B}}};$$

$$2. \ B_1^+ = \mathbf{1}_{d_{\mathcal{B}}}, \ B_1^- = \mathbf{0}_{d_{\mathcal{B}} \times d_{\mathcal{B}}},$$

$$\mathcal{B}_{\text{CH}} = A_1^+ \otimes B_2^+ - A_2^- \otimes \mathbf{1}_{d_{\mathcal{B}}} - A_2^+ \otimes B_2^+ = -A_1^- \otimes B_2^+ - A_2^- \otimes B_2^- \leq \mathbf{0}_{d_{\mathcal{A}} d_{\mathcal{B}} \times d_{\mathcal{A}} d_{\mathcal{B}}};$$

$$3. \ B_2^+ = \mathbf{0}_{d_{\mathcal{B}} \times d_{\mathcal{B}}}, \ B_2^- = \mathbf{1}_{d_{\mathcal{B}}},$$

$$\mathcal{B}_{\text{CH}} = -A_1^+ \otimes B_1^- - A_2^- \otimes B_1^+ \leq \mathbf{0}_{d_{\mathcal{A}} d_{\mathcal{B}} \times d_{\mathcal{A}} d_{\mathcal{B}}};$$

²This necessarily implies that the complementary POVM element is a zero matrix.

$$4. B_2^+ = \mathbf{1}_{d_B}, B_2^- = \mathbf{0}_{d_B \times d_B},$$

$$\mathcal{B}_{\text{CH}} = A_1^+ \otimes B_1^+ - A_2^- \otimes B_1^+ - A_2^+ \otimes \mathbf{1}_{d_B} = -A_1^- \otimes B_1^+ - A_2^+ \otimes B_1^- \leq \mathbf{0}_{d_A d_B \times d_A d_B}.$$

Since the Bell-CH inequality is symmetrical with respect to swapping the two parties, exactly the same argument can be applied to show that the resulting Bell-CH operator is NSD if any of Alice's POVM element is a full rank projector. Hence, the Bell-CH inequality cannot be violated in a standard Bell experiment involving at most two possible outcomes and where one of the measurement devices always gives the same measurement outcome.

B.3.2 Derivation of Horodecki's Criterion using LB

In order to determine if a general two-qubit state ρ violates the Bell-CH inequality, Eq. (3.23), we will have to first obtain an explicit expression for $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$. This can be done, for example, by evaluating Eq. (5.15) which, in turn, requires us to know the eigenvalues of $\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-}$, Eq. (5.13) for all s_b .

In this regard, let us note that for the Bell-CH inequality, we always have

$$\rho_{B_1^+} - \rho_{B_1^-} = \text{tr}_{\mathcal{A}} \left\{ \rho \left[(A_1^+ - A_2^-) \otimes \mathbf{1}_{d_B} \right] \right\}, \quad (\text{B.10a})$$

$$\rho_{B_2^+} - \rho_{B_2^-} = \text{tr}_{\mathcal{A}} \left\{ \rho \left[(A_1^+ - A_2^+) \otimes \mathbf{1}_{d_B} \right] \right\}, \quad (\text{B.10b})$$

$$\sum_{s_b, o_b} \text{tr} \left(\rho_{B_{s_b}^{o_b}} \right) = -1, \quad (\text{B.10c})$$

since $b_{12}^{++} = -b_{11}^{+-} = -b_{21}^{-+} = -b_{22}^{--} = 1$ while all the other $b_{s_a s_b}^{o_a o_b} = 0$ [c.f. Eq. (5.10) and Eq. (5.39)].

When Alice's choice of POVM is given by Eq. (5.40), it follows from Eq. (5.37) that the above expressions can be written more explicitly as

$$\begin{aligned} \rho_{B_1^+} - \rho_{B_1^-} &= \frac{1}{4} \left[\mathbf{r}_{\mathcal{A}} \cdot (\hat{a}_1 + \hat{a}_2) \mathbf{1}_2 + \sum_{i,j=x,y,z} (\hat{a}_1 + \hat{a}_2)_i [T]_{ij} \sigma_j \right], \\ \rho_{B_2^+} - \rho_{B_2^-} &= \frac{1}{4} \left[\mathbf{r}_{\mathcal{A}} \cdot (\hat{a}_1 - \hat{a}_2) \mathbf{1}_2 + \sum_{i,j=x,y,z} (\hat{a}_1 - \hat{a}_2)_i [T]_{ij} \sigma_j \right], \end{aligned}$$

which gives, respectively, eigenvalues

$$\lambda_1^{\pm} = \frac{1}{2} (\cos \theta \hat{c} \cdot \mathbf{r}_{\mathcal{A}} \pm |\cos \theta| ||T^T \hat{c}||), \quad \lambda_2^{\pm} = \frac{1}{2} (\sin \theta \hat{c}' \cdot \mathbf{r}_{\mathcal{A}} \pm |\sin \theta| ||T^T \hat{c}'||) \quad (\text{B.11})$$

where $\hat{c} \in \mathbb{R}^3$ and $\hat{c}' \in \mathbb{R}^3$ are orthogonal unit vectors defined via

$$\hat{a}_1 + \hat{a}_2 \equiv 2 \cos \theta \hat{c}, \quad \hat{a}_1 - \hat{a}_2 \equiv 2 \sin \theta \hat{c}'. \quad (\text{B.12})$$

We can now write Eq. (5.15) as

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, \hat{c}, \hat{c}', \theta) = \frac{1}{2} \sum_{s_b=1}^2 \sum_{o_b=\pm} |\lambda_{s_b}^{o_b}(\hat{c}, \hat{c}', \theta)| - \frac{1}{2}, \quad (\text{B.13})$$

which is to be maximized over all legitimate choices of \hat{c} , \hat{c}' and θ to give $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$.

Let us now consider the case in which $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$ is obtained by choosing \hat{c} , \hat{c}' and θ in $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, \hat{c}, \hat{c}', \theta)$ such that $\text{sgn}(\lambda_{s_b}^+) \neq \text{sgn}(\lambda_{s_b}^-)$ for *all* s_b . In this case, Eq. (B.13) becomes

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, \hat{c}, \hat{c}', \theta) = \frac{1}{2} (||T^T \hat{c}|| \cos \theta + ||T^T \hat{c}'|| \sin \theta) - \frac{1}{2}, \quad \theta \in \left[0, \frac{\pi}{4}\right], \quad (\text{B.14})$$

where we have redefined θ such that it now falls within 0 and $\pi/4$. The maximization over θ can now be carried out by choosing $\theta = \theta^*$ such that $||T^T \hat{c}|| \sin \theta^* = ||T^T \hat{c}'|| \cos \theta^*$, i.e.,³

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, \hat{c}, \hat{c}', \theta^*) = \frac{1}{2} \sqrt{||T^T \hat{c}||^2 + ||T^T \hat{c}'||^2} - \frac{1}{2}. \quad (\text{B.15})$$

From here, it suffices to choose \hat{c} and \hat{c}' as the (orthonormal) eigenvectors of $T T^T$ corresponding to the two largest eigenvalues. When arranged in descending order, the k^{th} eigenvalue of $T T^T$, however, is just the square of the k^{th} singular value of T , which we shall denote by ς_k . Hence, in this particular case, we have

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) = \frac{1}{2} \sqrt{\varsigma_1^2 + \varsigma_2^2} - \frac{1}{2}. \quad (\text{B.16})$$

What about the other cases in which $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho)$ is obtained by choosing \hat{c} , \hat{c}' and θ in $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho, \hat{c}, \hat{c}', \theta)$ such that $\text{sgn}(\lambda_{s_b}^+) = \text{sgn}(\lambda_{s_b}^-)$ for at least one of the s_b 's? In these cases, it follows from our discussion in Sec. 5.2.2.3 that for each of such s_b 's, the corresponding pair of optimal $B_{s_b}^{o_b}$ is given by $\{\mathbf{0}, \mathbf{1}_2\}$. However, as we have seen in Appendix B.3.1, if any of Alice's (or Bob's) POVM element is a full rank projector, the corresponding Bell operator is an NSD matrix, and hence cannot lead to a Bell-CH inequality violation. Moreover, the best that one can do in this case is to pick a classical strategy such that

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) = 0, \quad (\text{B.17})$$

which is necessarily greater than $\sqrt{\varsigma_1^2 + \varsigma_2^2}/2 - 1/2$.

Therefore, for a general two qubit state ρ , we have

$$\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) = \max \left\{ 0, \frac{1}{2} \sqrt{\varsigma_1^2 + \varsigma_2^2} - \frac{1}{2} \right\}. \quad (\text{B.18})$$

Recall that a Bell-CH violation by ρ is possible if and only if $\mathcal{S}_{\text{QM}}^{(\text{CH})}(\rho) > 0$. Hence, a two-qubit state violates the Bell-CH inequality if and only if

$$\varsigma_1^2 + \varsigma_2^2 > 1. \quad (5.42)$$

³That this choice of θ is always possible follows from the well-known fact that all singular values of T are less than or equal to one (see, for example, pp. 1840 of Ref. [217]).

B.4 Bell-Inequality Violations by Quantum States

B.4.1 Bell-CH-violation for Pure Two-Qudit States

Here, we will provide more details about the intermediate calculations leading to Eq. (6.3) and the corresponding optimal measurements, i.e., $\{B_{s_b}^\pm\}_{s_b=1}^2$, that should be carried out by Bob.

To begin with, we note from Eq. (6.2) that

$$A_1^+ - A_2^\mp = \frac{1}{2} (Z \pm X) = \frac{1}{2} \left[\bigoplus_{i=1}^{\lfloor d/2 \rfloor} X_\pm + (1 \pm 1) \Xi \right], \quad (\text{B.19})$$

where

$$X_\pm \equiv \sigma_z \pm \sigma_x = \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & -1 \end{pmatrix},$$

$$[\Xi]_{ij} = 0 \quad \forall \quad i, j \neq d, \quad [\Xi]_{dd} \equiv \xi = d \bmod 2;$$

here and below, whenever d is odd, we will assume that the end product of the direct sum is appended with zero entries to make the dimension of the resulting matrix $d \times d$.

From Eq. (6.1) and Eq. (B.10), it then follows that

$$\begin{aligned} \rho_{B_1^+} - \rho_{B_1^-} &= \frac{1}{2} \sum_{i,j=1}^{2\lfloor d/2 \rfloor} c_i c_j \mathcal{A}[j] \left(\bigoplus_{n=1}^{\lfloor d/2 \rfloor} X_+ \right) |i\rangle_{\mathcal{A}} |i\rangle_{\mathcal{B}\mathcal{B}} \langle j| + c_d^2 \Xi, \\ &= \frac{1}{2} \bigoplus_{n=1}^{\lfloor d/2 \rfloor} \begin{pmatrix} c_{2n-1}^2 & c_{2n-1} c_{2n} \\ c_{2n-1} c_{2n} & -c_{2n}^2 \end{pmatrix} + c_d^2 \Xi, \end{aligned} \quad (\text{B.20a})$$

and

$$\rho_{B_2^+} - \rho_{B_2^-} = \frac{1}{2} \bigoplus_{n=1}^{\lfloor d/2 \rfloor} \begin{pmatrix} c_{2n-1}^2 & -c_{2n-1} c_{2n} \\ -c_{2n-1} c_{2n} & -c_{2n}^2 \end{pmatrix}, \quad (\text{B.20b})$$

Some further calculations show that both these matrices have the following $2\lfloor d/2 \rfloor$ eigenvalues

$$\lambda_{n,\pm} = \frac{1}{4} (c_{2n-1}^2 - c_{2n}^2 \pm \kappa_n), \quad n = 1, 2, \dots, \left\lfloor \frac{d}{2} \right\rfloor,$$

where

$$\kappa_n \equiv \sqrt{(c_{2n-1}^2 + c_{2n}^2)^2 + 4c_{2n-1}^2 c_{2n}^2}.$$

For each n , let us denote the eigenvectors of $\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-}$ corresponding to eigenvalue $\lambda_{n,\pm}$ as $|v_{n,\pm}^{s_b}\rangle$, then, it can be shown that these eigenvectors only have the following nonzero entries

$$\begin{aligned} [|v_{n,\pm}^1\rangle]_{2n-1} &= \eta_{n,\mp} (c_{2n-1}^2 + c_{2n}^2 \pm \kappa_n), & [|v_{n,\pm}^1\rangle]_{2n} &= 2\eta_{n,\mp} c_{2n-1} c_{2n}, \\ [|v_{n,\pm}^2\rangle]_{2n-1} &= \eta_{n,\mp} (c_{2n-1}^2 + c_{2n}^2 \pm \kappa_n), & [|v_{n,\pm}^2\rangle]_{2n} &= -2\eta_{n,\mp} c_{2n-1} c_{2n}. \end{aligned}$$

where

$$\eta_{n,\pm} = \sqrt{\frac{\kappa_n \pm (c_{2n-1}^2 + c_{2n}^2)}{8 c_{2n-1}^2 c_{2n}^2 \kappa_n}}$$

is a normalization constant. When d is odd, $\rho_{B_1^+} - \rho_{B_1^-}$ and $\rho_{B_2^+} - \rho_{B_2^-}$, respectively, also have the eigenvalue c_d^2 and 0. In this case, the additional eigenvector of $\rho_{B_{s_b}^+} - \rho_{B_{s_b}^-}$, denoted by $|v_d^{s_b}\rangle$, where $s_b = 1, 2$, only has the following nonzero entry $[|v_d^{s_b}\rangle]_d = 1$.

Following the arguments presented in Sec. 5.2.2.3, we then know that the corresponding optimal measurements for Bob can be chosen to be

$$B_{s_b}^+ = \sum_{n=1}^{\lfloor d/2 \rfloor} |v_{n,+}^{s_b}\rangle\langle v_{n,+}^{s_b}| + \xi |v_d^{s_b}\rangle\langle v_d^{s_b}|, \quad B_{s_b}^- = \mathbb{1}_{d_B} - B_{s_b}^+. \quad (\text{B.21})$$

Moreover, the corresponding expectation value of Bell operator reads

$$\begin{aligned} \langle \mathcal{B}_{\text{CH}} \rangle_{|\Phi_d\rangle} &= \frac{1}{2} \sum_{s_b=1}^2 \left\| \rho_{B_{s_b}^+} - \rho_{B_{s_b}^-} \right\| + \frac{1}{2} \sum_{s_b} \sum_{o_b=\pm} \text{tr} \left(\rho_{B_{s_b}^{o_b}} \right), \\ &= \frac{1}{4} \sum_{s_b=1}^2 \sum_{n=1}^{\lfloor d/2 \rfloor} \kappa_n + \frac{\xi}{2} c_d^2 - \frac{1}{2}, \\ &= \frac{1}{2} \sum_{n=1}^{\lfloor d/2 \rfloor} \sqrt{(c_{2n-1}^2 + c_{2n}^2)^2 + 4c_{2n}^2 c_{2n-1}^2} + \frac{\xi}{2} c_d^2 - \frac{1}{2}, \end{aligned}$$

where we have also made use of Eq. (B.10c) and the fact that

$$c_{2n-1}^2 - c_{2n}^2 < \sqrt{(c_{2n-1}^2 + c_{2n}^2)^2 + 4c_{2n-1}^2 c_{2n}^2}.$$

B.5 Nonstandard Bell Experiments and Hidden Nonlocality

B.5.1 Proof of Lemma 17

For ease of reference, let us reproduce Lemma 17 as follows:

Lemma 17. *Let $\Omega_x : [\mathbb{C}^2] \otimes [\mathbb{C}^2] \rightarrow [\mathcal{H}_A \otimes \mathbb{C}^2] \otimes [\mathcal{H}_B \otimes \mathbb{C}^2]$ be a family of maps, separable with respect to the partition denoted by the brackets. Let μ be a unit-trace, PSD matrix acting on $[\mathcal{H}_A] \otimes [\mathcal{H}_B]$ such that*

$$\mu^T \otimes H_{\frac{\pi}{4}} - \int dx \, \Omega_x(H_{\theta_x}) \geq 0, \quad (7.27)$$

where H_θ is defined in Eq. (7.15), then μ has to be separable.

In order to prove this Lemma, and therefore Theorem 16, it is necessary to use the constraint that the maps Ω_x are separable. The problem of characterizing the separable maps is hard in

general since it maps onto the separability problem for bipartite states. However it turns out only to be necessary to determine the set of separable maps that take Bell diagonal states to Bell diagonal states and this can be done exactly (Appendix A). In what follows, we will provide the details for the proof of this Lemma.

Proof. The proof basically consists of three main steps. Firstly, we will need to characterize the set of separable maps Ω_x that is relevant to Eq. (7.27). Then, we will need to determine the values of θ_x that are allowed by the matrix inequality. Once we have characterized the set of separable maps Ω_x and inputs H_{θ_x} that satisfy the matrix inequality (7.27), it can further be shown that μ is the result of a separable map acting on a separable state, and hence separable.

Now, let us begin by characterizing the set of separable maps $\Omega_x : [\mathbb{C}^2] \otimes [\mathbb{C}^2] \rightarrow [\mathcal{H}_A \otimes \mathbb{C}^2] \otimes [\mathcal{H}_B \otimes \mathbb{C}^2]$ that satisfy the matrix inequality (7.27). For future reference, we will also refer to the first and second (output) qubit space involved in Ω_x as $\mathcal{H}_{A''}$ and $\mathcal{H}_{B''}$ respectively. Now, recall that the Bell basis is defined as

$$|\Phi_1\rangle \equiv \frac{1}{\sqrt{2}} (|0\rangle|0\rangle \pm |1\rangle|1\rangle), \quad (\text{A.1a})$$

$$|\Phi_3\rangle \equiv \frac{1}{\sqrt{2}} (|0\rangle|1\rangle \pm |1\rangle|0\rangle). \quad (\text{A.1b})$$

It is easy to show that the matrices H_θ defined in Eq. (7.15) are diagonal in this basis, i.e.,

$$H_\theta = \sum_{i=1}^4 [N_\theta]_i \Pi_i, \quad (\text{B.23a})$$

where $\Pi_i \equiv |\Phi_i\rangle\langle\Phi_i|$ ($i = 1, 2, 3, 4$) are the Bell projectors and $[N_\theta]_i$ is the i^{th} components of the vector

$$N_\theta \equiv \begin{pmatrix} 1 - \cos \theta - \sin \theta \\ 1 + \cos \theta - \sin \theta \\ 1 - \cos \theta + \sin \theta \\ 1 + \cos \theta + \sin \theta \end{pmatrix}. \quad (\text{B.23b})$$

For each value of x , let us now define the sixteen matrices

$$\omega_x^{ij} \equiv \text{tr}_{A''B''}[(\mathbf{1} \otimes \Pi_i) \Omega_x(\Pi_j)], \quad i, j = 1, 2, 3, 4, \quad (\text{B.24})$$

where the identity matrix $\mathbf{1}$ acts on $\mathcal{H}_A \otimes \mathcal{H}_B$ and Π_i acts on $\mathcal{H}_{A''} \otimes \mathcal{H}_{B''}$. Each ω_x^{ij} is the result of a physical operation, and hence PSD. Projecting the *lhs* of the matrix inequality (7.27) using the four Bell projectors Π_i , and taking the partial trace over $\mathcal{H}_{A''} \otimes \mathcal{H}_{B''}$, we get

$$\mu^T [N_{\frac{x}{4}}]_i - \int dx \sum_{j=1}^4 \omega_x^{ij} [N_{\theta_x}]_j \geq 0, \quad i = 1, 2, 3, 4. \quad (\text{B.25})$$

We shall also define a 4×4 matrix M_x whose (i, j) component is given by the trace of the corresponding ω_x^{ij} , i.e.,⁴

$$[M_x]_{i,j} \equiv \text{tr} \omega_x^{ij}. \quad (\text{B.26})$$

⁴Any of these 4×4 matrices is essentially the Jamiolkowski state corresponding to a separable, Bell-diagonal-preserving map written in the tensored Bell basis (Appendix A).

Performing the trace on the *lhs* of the matrix inequality (B.25), we obtain four inequalities which are associated with each of the four components of $N_{\frac{\pi}{4}}$,

$$N_{\frac{\pi}{4}} - \int dx M_x \cdot N_{\theta_x} \succeq \mathbf{0}_4, \quad (\text{B.27})$$

where $\mathbf{0}_4$ is the 4-dimensional null vector, and the symbols \cdot and \succeq mean, respectively, standard matrix multiplication and component-wise inequality.

Consider the set of matrices M that are generated by tracing the *lhs* of Eq. (B.24) when $\Omega_x : [\mathbb{C}^2] \otimes [\mathbb{C}^2] \rightarrow [\mathcal{H}_A \otimes \mathbb{C}^2] \otimes [\mathcal{H}_B \otimes \mathbb{C}^2]$ is any separable map. The characterization of this set of matrices can be found in Appendix A.1. In particular, let us denote by \mathcal{D} and \mathcal{G} , respectively, the convex hull of all matrices obtained by independently permuting the rows and/or columns of D_0 and G_0 , c.f. Eq. (A.5). It then follows from Definition 18, Theorem 19 and Choi-Jamiołkowski isomorphism (Appendix A.3) that any matrix M as defined above can be written as

$$M = p D + q G, \quad (\text{B.28})$$

where $D \in \mathcal{D}$, $G \in \mathcal{G}$, and $p, q \geq 0$. Then, any solution to the vector inequality (B.27) can be labeled by giving $(\theta_x, p_x, q_x, D_x, G_x)$.

Now, let us characterize the set of admissible solutions to the vector inequality (B.27). By using the fact that $G \cdot N_\theta \succeq \mathbf{0}_4$ for all θ and all $G \in \mathcal{G}$, we can see that any solution of the vector inequality (B.27) must satisfy

$$N_{\frac{\pi}{4}} \succeq \int dx p_x D_x \cdot N_{\theta_x}. \quad (\text{B.29})$$

Recall that this component-wise inequality entails four inequalities. Adding them together we obtain the condition

$$\int dx p_x \leq 4. \quad (\text{B.30})$$

Denote by \mathcal{N} the set of all vectors obtained by permuting the components of N_θ , Eq. (B.23b), when θ runs through $[0, \pi/4]$. With some thought, it is not difficult to see that the convex hull of \mathcal{N} , denoted by $\text{conv}(\mathcal{N})$, is precisely the set of vectors that can be written as the *rhs* of the vector inequality (B.29) under the constraint given by Eq. (B.30). We can then write the first inequality of (B.29) as

$$1 - \sqrt{2} \geq [N]_1, \quad (\text{B.31})$$

where $[N]_1$ is the first component of $N \in \text{conv}(\mathcal{N})$. It is easy to see that all vectors $N \in \text{conv}(\mathcal{N})$ satisfy the converse inequality, namely, $1 - \sqrt{2} \leq [N]_1$, and only $N_{\frac{\pi}{4}}$ saturates it. Hence, the only admissible solution for the *rhs* of the vector inequality (B.29) is $N_{\frac{\pi}{4}}$. Substituting this into the vector inequality (B.27), and again using Eq. (B.28) and Eq. (B.30), we obtain $-\int dx q_x G_x \cdot N_{\theta_x} \succeq \mathbf{0}_4$. However, as mentioned above, $G \cdot N_\theta \succeq \mathbf{0}_4$ for all θ and all $G \in \mathcal{G}$, which implies that for any solution to the vector inequality (B.27), we must have $\int dx q_x G_x \cdot N_{\theta_x} = \mathbf{0}_4$. Therefore, the vector inequality (B.27) may now be written as

$$N_{\frac{\pi}{4}} - M_0 \cdot N_{\frac{\pi}{4}} \succeq \mathbf{0}_4, \quad (\text{B.32})$$

where M_0 is any doubly-stochastic matrix such that

$$M_0 \cdot N_{\frac{\pi}{4}} = N_{\frac{\pi}{4}}. \quad (\text{B.33})$$

With some thought, it can be shown that the form of $N_{\frac{\pi}{4}}$ demands that doubly-stochastic matrices that satisfy Eq. (B.33) must have the following form

$$M_0 = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 - \eta & \eta & \cdot \\ \cdot & \eta & 1 - \eta & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \quad (\text{B.34})$$

where $\eta \in [0, 1]$.

On the other hand, the vector inequality (B.32) and Eq. (B.33) together imply that the *lhs* of the former, and hence Eq. (B.27) is $\mathbf{0}_4$. Since the four inequalities in Eq. (B.27) were obtained by taking the trace of the matrix inequality (B.25), this further implies that the *lhs* of the matrix inequality (B.25) is traceless for all i . The only positive matrix with zero trace is the null matrix, therefore we must have

$$\mu^T[N_{\frac{\pi}{4}}]_i = \sum_{j=1}^4 \omega_0^{ij}[N_{\frac{\pi}{4}}]_j, \quad i = 1, 2, 3, 4, \quad (\text{B.35})$$

where ω_0 is any ω_x that gives rise to M_0 . By the same token, c.f. Eq. (B.26), the pairs (i, j) for which $[M_0]_{i,j} = 0$ must have originated from ω_0^{ij} which is a null matrix.

Finally, if we now add the equalities in Eq. (B.35) corresponding to $i = 2, 3$, it follows from the definition of ω_0^{ij} [Eq. (B.24)] that

$$2 \mu^T = \text{tr}_{\mathcal{A}''\mathcal{B}''} [(\mathbf{1} \otimes \Psi) \Omega_0(\Psi)], \quad (\text{B.36})$$

where $\Psi = \Pi_2 + \Pi_3$, and Ω_0 is any Ω_x that gives rise to ω_0 . From the PPT criterion of separability [161, 162], one can easily check that the (unnormalized) two-qubit state Ψ is a separable state. Eq. (B.36) implies that μ^T is the output of a separable map applied to a separable input state, and hence is a separable state as we have wanted to prove. \square



Semidefinite Programming and Relaxations

C.1 Semidefinite Programs

A semidefinite program (SDP) is a convex optimization over Hermitian matrices [63, 64]. The objective function depends linearly on the matrix variable (as expectation values do in quantum mechanics for example) and the optimization is carried out subjected to the constraint that the matrix variable is positive semidefinite (PSD) and satisfies various affine constraints. Any semidefinite program may be written in the following *standard form*:

$$\text{maximize} \quad -\text{tr}[F_0 Z], \tag{C.1a}$$

$$\text{subject to} \quad \text{tr}[F_i Z] = c_i \quad \forall i, \tag{C.1b}$$

$$Z \geq 0, \tag{C.1c}$$

where F_0 and all the F_i 's are Hermitian matrices and the c_i are real numbers that together specify the optimization; Z is the Hermitian matrix variable to be optimized.

An SDP also arises naturally in the *inequality form*, which seeks to minimize a linear function of the optimization variables $\mathbf{x} \in \mathbb{R}^n$, subjected to a linear matrix inequality (LMI):

$$\text{minimize} \quad \mathbf{x}^T \mathbf{c}' \tag{C.2a}$$

$$\text{subject to} \quad G_0 + \sum_i [\mathbf{x}]_i G_i \geq 0. \tag{C.2b}$$

As in the standard form, G_0 and all the G_i 's are Hermitian matrices, while \mathbf{c}' is a real vector of length n .

C.2 Semidefinite Relaxation to Finding $\mathcal{S}_{\text{QM}}(\rho)$

The global optimization problem of finding $\mathcal{S}_{\text{QM}}(\rho)$, either in the form of Eq. (5.18) for a two-outcome Bell correlation inequality, or Eq. (5.44) for a two-outcome Bell inequality for probabilities, is a QCQP. As was demonstrated in Sec. 5.2.3, an upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ can then be obtained by considering the corresponding *Lagrange Dual*.

More generally, the global optimization problem of finding $\mathcal{S}_{\text{QM}}(\rho)$ can be mapped to a real polynomial optimization problem:

$$\text{maximize } f_{\text{obj}}(\mathbf{y}), \quad (\text{C.3a})$$

$$\text{subject to } f_{\text{eq},i}(\mathbf{y}) = 0, \quad i = 1, 2, \dots, N_{\text{eq}}, \quad (\text{C.3b})$$

$$f_{\text{ineq},j}(\mathbf{y}) \geq 0, \quad j = 1, 2, \dots, N_{\text{ineq}}, \quad (\text{C.3c})$$

where \mathbf{y} is a vector of *real* variables formed by the expansion coefficients of local observables $\{O_m\}$ in terms of Hermitian basis operators.

By considering Positivstellensatz-based relaxations, a hierarchy of upper bounds for $f_{\text{obj}}(\mathbf{y})$ can be obtained by solving appropriate SDPs (see, for example, Ref. [183] and references therein). To see this, let us first note that γ will be an upper bound on the constrained optimization problem (C.3) if there exists a set of sum of squares (SOS) $\mu_i(\mathbf{y})$'s (i.e., non-negative, real polynomials that can be written as $\sum_j [h_j(\mathbf{y})]^2$ with $h_j(\mathbf{y})$ being some real polynomials of \mathbf{y}), and a set of real polynomials $\nu_j(\mathbf{y})$ such that [183, 191, 192]

$$\begin{aligned} \gamma - f_{\text{obj}}(\mathbf{y}) = & \mu_0(\mathbf{y}) + \sum_j \nu_j(\mathbf{y}) f_{\text{eq},j}(\mathbf{y}) + \sum_i \mu_i(\mathbf{y}) f_{\text{ineq},i}(\mathbf{y}) \\ & + \sum_{i_1, i_2} \mu_{i_1, i_2}(\mathbf{y}) f_{\text{ineq},i_1}(\mathbf{y}) f_{\text{ineq},i_2}(\mathbf{y}) + \dots \end{aligned} \quad (\text{C.4})$$

The relaxed optimization problem then consists of minimizing γ subjected to the above constraint. Clearly, at values of \mathbf{y} where the constraints are satisfied, γ gives an upper bound on $f_{\text{obj}}(\mathbf{y})$. The auxiliary polynomials $\nu_j(\mathbf{y})$ and SOS $\mu_i(\mathbf{y})$ are thus analogous to the Lagrange multipliers in the relaxed optimization problem.

For a fixed degree of the above expression, this relaxed optimization problem can be cast as an SDP in the form of Eq. (C.2) [183]. For the lowest order relaxation, the auxiliary polynomials $\nu_j(\mathbf{y})$ and SOS $\mu_i(\mathbf{x})$ are chosen such that degree of the expression in Eq. (C.4) is no larger than the maximum degree of the set of polynomials

$$f_{\text{obj}}(\mathbf{y}), f_{\text{eq},1}(\mathbf{y}), \dots, f_{\text{eq},N_{\text{eq}}}(\mathbf{y}), f_{\text{ineq},1}(\mathbf{y}), \dots, f_{\text{ineq},N_{\text{ineq}}}(\mathbf{y});$$

for a QCQP with no inequality constraints, this amounts to setting all the $\mu_i(\mathbf{y})$ to zero and all the $\nu_j(\mathbf{y})$ to numbers.

For higher order relaxation, we increase the degree of the expression in Eq. (C.4) by increasing the degree of the auxiliary polynomials. At the expense of involving more computational resources, a tighter upper bound on $f_{\text{obj}}(\mathbf{y})$ can then be obtained by solving the corresponding SDP.

C.2.1 Lowest Order Relaxation with Observables of Fixed Trace

We have seen in Sec. 5.2.3.2 that a direct implementation of the Lagrange dual to the optimization problem given in Eq. (5.18) — disregarding the constraint given by (5.31) — gives rise to an upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ that is apparently state-independent. To obtain a tighter upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ using again the lowest order relaxation to Eq. (5.18), we found it most convenient to express the optimization problem in terms of the *real* optimization variables,

$$y_{mn} \equiv \text{tr}(O_m \sigma_n), \quad n = 0, 1, \dots, d^2 - 1, \quad (\text{C.5})$$

which are just the expansion coefficients of each O_m in terms of a set of Hermitian basis operators $\{\sigma_n\}_{n=0}^{d^2-1}$ satisfying Eq. (C.15). The constraint (5.31) can then be taken care of by setting each $y_{m0} = z_m/\sqrt{d}$. It is also expedient to express the density matrix ρ in terms of the same basis of Hermitian operators

$$\rho = \frac{\mathbf{1}_d \otimes \mathbf{1}_d}{d^2} + \sum_{i=1}^{d^2-1} ([\mathbf{r}_{\mathcal{A}}]_i \sigma_i \otimes \sigma_0 + [\mathbf{r}_{\mathcal{B}}]_i \sigma_0 \otimes \sigma_i) + \sum_{i,j=1}^{d^2-1} [R']_{ij} \sigma_i \otimes \sigma_j \quad (\text{C.6a})$$

where

$$[R']_{ij} = \text{tr}(\rho \sigma_i \otimes \sigma_j), \quad (\text{C.6b})$$

$$[\mathbf{r}_{\mathcal{A}}]_i \equiv \text{tr}(\rho \sigma_i \otimes \sigma_0), \quad [\mathbf{r}_{\mathcal{B}}]_j \equiv \text{tr}(\rho \sigma_0 \otimes \sigma_j); \quad (\text{C.6c})$$

$\mathbf{r}_{\mathcal{A}}$, $\mathbf{r}_{\mathcal{B}}$ are simply the coherence vectors that have been studied in the literature [193, 194].

We will now incorporate the constraints (5.31) by expressing the Lagrangian (5.23) as a function of the reduced set of variables

$$(\mathbf{y}')^T \equiv [y_{11} \ y_{12} \ \dots \ y_{1 \ d^2-1} \ y_{21} \ \dots \ y_{m_{\mathcal{A}}+m_{\mathcal{B}} \ d^2-1}], \quad (\text{C.7})$$

while all the $y_{m0} = z_m/\sqrt{d}$ are treated as fixed parameters of the problem. With this change in basis, and after some patient algebra, the Lagrangian can be rewritten as

$$\begin{aligned} \mathcal{L}(\mathbf{y}', \lambda_{mn}) = & \sum_{m=1}^{m_{\mathcal{A}}+m_{\mathcal{B}}} \lambda_{m0} \left(\sqrt{d} - \frac{z_m^2}{d\sqrt{d}} \right) + \sum_{s_{\mathcal{A}}=1}^{m_{\mathcal{A}}} \sum_{s_{\mathcal{B}}=1}^{m_{\mathcal{B}}} b_{s_{\mathcal{A}}s_{\mathcal{B}}} \frac{z_{s_{\mathcal{A}}} z_{s_{\mathcal{B}}+m_{\mathcal{A}}}}{d^2} \\ & - \frac{1}{\sqrt{d}} (\mathbf{l} - \mathbf{r})^T (\mathbf{y}') - (\mathbf{y}')^T \Omega' (\mathbf{y}'), \end{aligned} \quad (\text{C.8})$$

where λ_{mn} are defined in Eq. (5.30),

$$\begin{aligned} \Omega' \equiv & \frac{1}{2} \begin{pmatrix} \mathbf{0}_{m_{\mathcal{A}}(d^2-1) \times m_{\mathcal{A}}(d^2-1)} & -b \otimes R' \\ -(b \otimes R')^T & \mathbf{0}_{m_{\mathcal{B}}(d^2-1) \times m_{\mathcal{B}}(d^2-1)} \end{pmatrix} + \bigoplus_{m=1}^{m_{\mathcal{A}}+m_{\mathcal{B}}} M_m, \\ \mathbf{l} \equiv & \text{vec}(L), \quad \mathbf{r} \equiv \begin{pmatrix} \mathbf{t}_{\mathcal{A}} \otimes \mathbf{r}_{\mathcal{A}} \\ \mathbf{t}_{\mathcal{B}} \otimes \mathbf{r}_{\mathcal{B}} \end{pmatrix}, \end{aligned} \quad (\text{C.9a})$$

and for $i, j = 1, 2, \dots, d^2 - 1$,

$$\begin{aligned} [L]_{j,m} &= 2z_m \lambda_{mj}, \quad [\mathbf{t}_A]_{s_a} = \sum_{s_b=1}^{m_B} b_{s_a s_b} z_{s_b+m_A}, \quad [\mathbf{t}_B]_l = \sum_{s_a=1}^{m_A} b_{s_a s_b} z_{s_a}, \\ M_m &= \sum_{n=0}^{d^2-1} \lambda_{mn} P_n, \quad [P_n]_{i,j} = \frac{1}{2} \text{tr} (\sigma_n [\sigma_i, \sigma_j]_+); \end{aligned} \quad (\text{C.9b})$$

$[\sigma_i, \sigma_j]_+ \equiv \sigma_i \sigma_j + \sigma_j \sigma_i$ is the anti-commutator of σ_i and σ_j .

As before, we now maximize the Lagrangian (C.8) over \mathbf{y}' to obtain the corresponding Lagrange dual function. The latter, however, is unbounded above unless

$$\begin{pmatrix} -2t & \frac{1}{\sqrt{d}} (\mathbf{1}^T - \mathbf{r}^T) \\ \frac{1}{\sqrt{d}} (\mathbf{1} - \mathbf{r}) & 2\Omega' \end{pmatrix} \geq 0, \quad (\text{C.10})$$

for some finite t . The convex optimization problem dual to Eq. (5.18) with fixed trace for each observables is thus

$$\begin{aligned} \text{minimize} \quad & \sum_{m=1}^{m_A+m_B} \lambda_{m0} \left(\sqrt{d} - \frac{z_m^2}{d\sqrt{d}} \right) + \sum_{s_a=1}^{m_A} \sum_{s_b=1}^{m_B} b_{s_a s_b} \frac{z_{s_a} z_{s_b+m_A}}{d^2} - t, \\ \text{subject to} \quad & \begin{pmatrix} -2t & \frac{1}{\sqrt{d}} (\mathbf{1}^T - \mathbf{r}^T) \\ \frac{1}{\sqrt{d}} (\mathbf{1} - \mathbf{r}) & 2\Omega' \end{pmatrix} \geq 0. \end{aligned} \quad (\text{C.11})$$

C.2.2 Sufficient Condition for No-violation of the Bell-CHSH Inequality

To derive the semianalytic criterion Eq. (5.35), we now note that any choice of $\{\lambda_{mn}\}_{n=0}^{d^2-1}$ that satisfy constraint (C.10) will provide an upper bound on the corresponding $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$. In particular, an upper bound can be obtained by setting

$$\lambda_{mn} = \delta_{n0} [\lambda_A (\delta_{m1} + \delta_{m2}) + \lambda_B (\delta_{m3} + \delta_{m4})], \quad (\text{C.12})$$

and solving for λ_A, λ_B that satisfy the constraint (C.10). With this choice of the Lagrange multipliers, and for quantum states with vanishing coherence vectors, the constraint (C.10) becomes

$$\begin{pmatrix} -2t & \mathbf{0}_{2(d^2-1)}^T & \mathbf{0}_{2(d^2-1)}^T \\ \mathbf{0}_{2(d^2-1)} & \frac{2\lambda_A}{\sqrt{d}} \mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} & -b \otimes R' \\ \mathbf{0}_{2(d^2-1)} & -(b \otimes R')^T & \frac{2\lambda_B}{\sqrt{d}} \mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} \end{pmatrix} \geq 0, \quad (\text{C.13})$$

where b and R' are defined, respectively, in Eq. (5.33) and Eq. (C.6b). This, in turn is equivalent to

$$-t \geq 0, \quad (\text{C.14a})$$

$$\begin{pmatrix} \frac{2\lambda_A}{\sqrt{d}} \mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} & -b \otimes R' \\ -(b \otimes R')^T & \frac{2\lambda_B}{\sqrt{d}} \mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} \end{pmatrix} \geq 0. \quad (\text{C.14b})$$

Using Schur's complement [186, 187] and Eq. (5.33), the constraint (C.14b) can be explicitly solved to give

$$\lambda_{\mathcal{A}}\lambda_{\mathcal{B}} \geq \frac{1}{2}s_1^2d,$$

where s_1 is the largest singular value of the matrix R' . Substituting this and Eq. (C.14a) into Eq. (C.11), and after some algebra, we see that $\mathcal{S}_{\text{QM}}^{(\text{CHSH})}(\rho)$ for a quantum state ρ with vanishing coherence vectors cannot be greater than

$$\max_{z_1, z_2, z_3, z_4} 2\sqrt{2}s_1d \sqrt{\prod_{i=1}^2 \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{s_a, s_b=1}^2 b_{s_a s_b} \frac{z_{s_a} z_{s_b+2}}{d^2}.$$

For ρ to violate the Bell-CHSH inequality, we must have this upper bound greater than the classical threshold value, $\beta_{\text{LHV}}^{(\text{CHSH})} = 2$, c.f. Eq. (5.5). Hence a sufficient condition for ρ to satisfy the Bell-CHSH inequality is given by Eq. (5.35).

C.3 Explicit Forms of Semidefinite Programs

C.3.1 SDP for the LB Algorithm

Here, we provide an explicit form for the matrices F_i and constants c_i that define the SDP used in the LB algorithm, Eq. (5.14). By setting

$$Z = \begin{pmatrix} B_1^1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & B_1^{n_B} & 0 & 0 & 0 \\ 0 & 0 & 0 & B_2^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{m_B}^{n_B} \end{pmatrix},$$

in Eq. (C.1), we see that the inequality constraint (C.1c) of the SDP entails the positive semidefiniteness of the POVM elements $\left\{ \{B_{s_b}^{o_b}\}_{o_b=1}^{n_B} \right\}_{s_b=1}^{m_B}$, and hence Eq. (5.14c). On the other hand, with

$$F_0 = - \begin{pmatrix} \rho_{B_1^1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_{B_1^{n_B}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_{B_2^1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_{B_{m_B}^{n_B}} \end{pmatrix},$$

where $\rho_{B_{s_b}^{o_b}}$ is defined in Eq. (5.13), the equality constraint (C.1b), together with appropriate choice of F_i and c_i , ensures that the normalization condition (5.14b) is satisfied.

In particular, each F_i is formed from a direct sum of Hermitian basis operators. A convenient choice of such basis operators is given by the traceless Gell-Mann matrices, denoted by $\{\sigma_n\}_{n=1}^{d^2-1}$, supplemented by

$$\sigma_0 = \frac{1}{\sqrt{d}} \mathbf{1}_d, \quad (\text{C.15a})$$

such that

$$\text{tr}(\sigma_n \sigma_{n'}) = \delta_{nn'} \quad \text{and} \quad \text{tr}(\sigma_n) = \sqrt{d} \delta_{n0}, \quad (\text{C.15b})$$

where $d = d_{\mathcal{B}}$ is the dimension of the state space that each $B_{s_b}^{o_b}$ acts on. A typical F_i then consists of $n_{\mathcal{B}}$ diagonal blocks of σ_n at positions corresponding to the $n_{\mathcal{B}}$ POVM elements $\{B_{s_b}^{o_b}\}_{o_b=1}^{n_{\mathcal{B}}}$ in Z for a fixed s_b . For instance, the set of F_i

$$F_i = \begin{pmatrix} \sigma_{i-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{i-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad 1 \leq i \leq d^2,$$

together with $c_i = \sqrt{d} \delta_{i1}$ entails the normalization of $\{B_{s_b}^{o_b}\}_{o_b=1}^{n_{\mathcal{B}}}$, i.e., $\sum_{o_b=1}^{n_{\mathcal{B}}} B_1^{o_b} = \mathbf{1}_{d_{\mathcal{B}}}$; the remaining $(m_{\mathcal{B}} - 1)d^2$ F_i are defined similarly and can be obtained by shifting the nonzero diagonal blocks diagonally downward by appropriate multiples of $n_{\mathcal{B}}$ blocks. The SDP thus consists of solving Eq. (C.1) for a $m_{\mathcal{B}} n_{\mathcal{B}} d \times m_{\mathcal{B}} n_{\mathcal{B}} d$ Hermitian matrix Z subjected to $d^2 m_{\mathcal{B}}$ affine constraints.

C.3.2 SDP for the UB Algorithm

In analogy with the previous section, we will provide, in this section, an explicit form for some of the SDPs used in the UB algorithm. In particular, we find it expedient to express these SDPs in the inequality form, Eq. (C.2), but for convenience, we will use two indices m and n ($m = 1, 2, \dots, m_{\mathcal{A}} + m_{\mathcal{B}}$, $n = 0, 1, \dots, d^2 - 1$), instead of the single index i [c.f. Eq. (C.2b)] to label the Hermitian matrices G_{mn} and the components of the vector \mathbf{c}' . Throughout this section, σ_n will refer to a Hermitian basis operator satisfying Eq. (C.15).

C.3.2.1 State-independent Bound

Now, we will give the matrices G_{mn} and constants $[\mathbf{c}']_{mn}$ that define the SDP obtained from the lowest order relaxation to Eq. (5.18) given by Eq. (5.29) and Eq. (5.30). To begin with, it is straightforward to see that by setting

$$[\mathbf{x}]_{mn} = \lambda_{mn}, \quad [\mathbf{c}']_{mn} = \sqrt{d} \delta_{n0}$$

in Eq. (C.2), we obtain the same objective function as that in Eq. (5.29), where λ_{mn} is defined in Eq. (5.30). Next, if we further set [c.f. Eq. (5.21) and Eq. (5.27)]

$$G_0 = \Omega_0 = \frac{1}{2} \begin{pmatrix} \mathbf{0}_{d^2 m_{\mathcal{A}} \times d^2 m_{\mathcal{A}}} & -b \otimes R \\ -b^T \otimes R^\dagger & \mathbf{0}_{d^2 m_{\mathcal{B}} \times d^2 m_{\mathcal{B}}} \end{pmatrix},$$

where b and R are defined just after Eq. (5.21), and

$$G_{mn} = \bigoplus_{k=1}^{m-1} \mathbf{0}_{d^2 \times d^2} \bigoplus \mathbf{1}_d \otimes \sigma_n \bigoplus_{k=m+1}^{m_A+m_B} \mathbf{0}_{d^2 \times d^2},$$

then it can be seen that Eq. (C.2b) enforces the inequality constraint given in Eq. (5.29). The SDP corresponding to Eq. (5.29), which apparently gives rise to a state-independent upper bound on $\mathcal{S}_{\text{QM}}(\rho)$, thus consists of solving Eq. (C.2) for $d^2(m_A + m_B)$ real variables subjected to a matrix inequality of dimension $d^2(m_A + m_B) \times d^2(m_A + m_B)$, and which is linear in the $d^2(m_A + m_B)$ real variables.

C.3.2.2 State-dependent Bound

For the more refined SDP given by Eq. (C.11), which gives rise to a state-dependent upper bound on $\mathcal{S}_{\text{QM}}(\rho)$, we will instead set

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_0 \oplus t, \quad \mathbf{c}' = \mathbf{c}_0 \oplus -1, \\ [\mathbf{x}_0]_{mn} &= \lambda_{mn}, \quad [\mathbf{c}_0]_{mn} = \left(\sqrt{d} - \frac{z_m^2}{d\sqrt{d}} \right) \delta_{n0}, \end{aligned}$$

in Eq. (C.2), where z_m is the trace of local observables defined in Eq. (5.31). It is easy to see that with the above choice of \mathbf{x} and \mathbf{c}' , Eq. (C.2a) gives, apart from a constant that is immaterial to the optimization, the same objective function as that in Eq. (C.11). Next, we will set

$$\begin{aligned} G_0 &= - \begin{pmatrix} 0 & \frac{1}{\sqrt{d}} (\mathbf{t}_A \otimes \mathbf{r}_A)^T & \frac{1}{\sqrt{d}} (\mathbf{t}_B \otimes \mathbf{r}_B)^T \\ \frac{1}{\sqrt{d}} \mathbf{t}_A \otimes \mathbf{r}_A & \mathbf{0}_{(d^2-1)m_A \times (d^2-1)m_A} & b \otimes R' \\ \frac{1}{\sqrt{d}} \mathbf{t}_B \otimes \mathbf{r}_B & (b \otimes R')^T & \mathbf{0}_{(d^2-1)m_B \times (d^2-1)m_B} \end{pmatrix}, \\ G_{mn} &= \left[0 \bigoplus_{k=1}^{m-1} \mathbf{0}_{(d^2-1) \times (d^2-1)} \bigoplus P_n \bigoplus_{k=m+1}^{m_A+m_B} \mathbf{0}_{(d^2-1) \times (d^2-1)} \right] + (1 - \delta_{n0}) \frac{2z_m}{\sqrt{d}} G'_{mn}, \end{aligned}$$

where $\mathbf{t}_A, \mathbf{t}_B, P_n$ are defined in Eq. (C.9), $\mathbf{r}_A, \mathbf{r}_B, R'$ are defined in Eq. (C.6) and G'_{mn} is a $[1 + (d^2 - 1)(m_A + m_B)] \times [1 + (d^2 - 1)(m_A + m_B)]$ matrix that is zero everywhere except for the following entries:

$$[G'_{mn}]_{1, 1+(m-1)(d^2-1)+n} = [G'_{mn}]_{1+(m-1)(d^2-1)+n, 1} = 1.$$

Finally, by setting

$$G_t = \begin{pmatrix} -2 & \mathbf{0}_{(d^2-1)(m_A+m_B)}^T \\ \mathbf{0}_{(d^2-1)(m_A+m_B)} & \mathbf{0}_{(d^2-1)(m_A+m_B), (d^2-1)(m_A+m_B)} \end{pmatrix},$$

in Eq. (C.2), which is the G_i corresponding to the variable t , it can be seen that Eq. (C.2b) enforces the matrix inequality constraint given in Eq. (C.11). The SDP corresponding to Eq. (C.11) thus consists of solving Eq. (C.2) for $d^2(m_A + m_B) + 1$ real variables subjected to a matrix inequality of dimension $[1 + (d^2 - 1)(m_A + m_B)] \times [1 + (d^2 - 1)(m_A + m_B)]$, and which is linear in the $d^2(m_A + m_B) + 1$ real variables.

C.3.3 SDP for the Verification of Entanglement Witness

Here, we will show that, in the context of Lemma 20, the search for a PSD \mathcal{Z} satisfying Eq. (A.10) is a semidefinite programming feasibility problem [63, 64], i.e., an SDP whereby the objective function is some constant that is independent of any optimization variables. In particular, we will show that this SDP is readily written in the *standard form*, Eq. (C.1), but for convenience, we will use two indices i and j instead of the single index i [c.f. Eq. (C.1b)] to label the Hermitian matrices F_{ij} and the constants c_{ij} .

Let us denote by $\{\sigma_i^A\}_{i=0}^{d_A'^2-1}$ and $\{\sigma_j^B\}_{j=0}^{d_B'^2-1}$, respectively, a complete set of Hermitian basis operators acting on $\Pi_A \mathcal{H}_A^{\otimes m} \Pi_A$ and $\Pi_B \mathcal{H}_B^{\otimes n} \Pi_B$ where Π_A, Π_B are, respectively, the projectors onto the symmetric subspace of $\mathcal{H}_A^{\otimes m}$ and $\mathcal{H}_B^{\otimes n}$ and d_A', d_B' are the corresponding dimensions of these symmetric subspaces. As before, a convenient choice of such basis operators is given by the orthonormal set which satisfies Eq. (C.15). Since both the *lhs* and *rhs* of Eq. (A.10) are Hermitian matrices, if the equation holds true, it follows that for all i and j we must have

$$\begin{aligned} \text{tr} [\Pi_A \otimes \Pi_B (\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \Pi_A \otimes \Pi_B \sigma_i^A \otimes \sigma_j^B] &= \text{tr} [\Pi_A \otimes \Pi_B \mathcal{Z}^{T_s} \Pi_A \otimes \Pi_B \sigma_i^A \otimes \sigma_j^B], \\ \Rightarrow \text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_i^A \otimes \sigma_j^B] &= \text{tr} [\mathcal{Z}^{T_s} \sigma_i^A \otimes \sigma_j^B], \\ \Rightarrow \text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_i^A \otimes \sigma_j^B] &= \text{tr} [\mathcal{Z} (\sigma_i^A \otimes \sigma_j^B)^{T_s}]. \end{aligned}$$

Moreover, it is easy to see that whenever this last expression holds true for all i and j , one can construct a PSD \mathcal{Z} such that Eq. (A.10) holds true. Hence, if we set

$$\begin{aligned} F_0 &= \mathbf{0}_{d_A^m d_B^n \times d_A^m d_B^n}, \quad F_{ij} = (\sigma_i^A \otimes \sigma_j^B)^{T_s}, \\ Z &= \mathcal{Z}, \quad c_{ij} = \text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_i^A \otimes \sigma_j^B], \end{aligned}$$

in Eq. (C.1), we will have expressed the problem of searching for a legitimate \mathcal{Z} as a semidefinite programming feasibility problem.

On the other hand, for numerical implementation of the above SDP, it may be advantageous to formalize the above problem as an ordinary SDP where F_0 is nonzero. For that purpose, one sets, instead,

$$\begin{aligned} F_0 &= (\sigma_0^A \otimes \sigma_0^B)^{T_s}, \quad F_{ij} = (\sigma_i^A \otimes \sigma_j^B)^{T_s}, \\ c_{ij} &= \text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_i^A \otimes \sigma_j^B], \end{aligned}$$

in Eq. (C.1), where now we have excluded F_{00} from the set of F_{ij} . With some thought, it is not difficult to see that a legitimate \mathcal{Z} that satisfies all the constraints exists if and only if the optimum of the optimization, Z^* satisfies

$$-\text{tr}(F_0 Z^*) \geq -\text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_0^A \otimes \sigma_0^B], \quad (\text{C.16})$$

in which case the desired \mathcal{Z} can be constructed as

$$\mathcal{Z} = Z^* + \{\text{tr} [(\mathbf{1}_{d_A}^{\otimes m-1} \otimes Z_w \otimes \mathbf{1}_{d_B}^{\otimes n-1}) \sigma_0^A \otimes \sigma_0^B] - \text{tr}(F_0 Z^*)\} \sigma_0^A \otimes \sigma_0^B. \quad (\text{C.17})$$

Hence, the search for a legitimate \mathcal{Z} can also be formalized as an SDP which consists of solving Eq. (C.1) for a $d_A^m d_B^n \times d_A^m d_B^n$ PSD matrix Z subjected to $(d_A'^2 - 1)(d_B'^2 - 1) - 1$ affine constraints.

List of Symbols

The following list is neither exhaustive nor exclusive, but may be helpful.

$p_{\mathcal{AB}}^{o_a o_b}(s_a, s_b)$ Joint probability that the o_a^{th} and o_b^{th} experimental outcomes are observed, respectively, at Alice's and Bob's site given that she performs the s_a^{th} and he performs the s_b^{th} measurement.

$p_{\mathcal{A}}^{o_a}(s_a)$ The marginal probability that the o_a^{th} experimental outcome is observed at Alice's site given that she performs the s_a^{th} measurement.

$p_{\mathcal{B}}^{o_b}(s_b)$ The marginal probability that the o_b^{th} experimental outcome is observed at Bob's site given that he performs the s_b^{th} measurement.

$\mathcal{C}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$ The set of probability vectors obeying the no-signaling conditions when Alice and Bob are allowed to perform, respectively, $m_{\mathcal{A}}$ and $m_{\mathcal{B}}$ alternative measurements and where each local measurement yields, correspondingly, one of $n_{\mathcal{A}}$ and $n_{\mathcal{B}}$ outcomes.

$\mathcal{P}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$ The set of classical probability vectors in $\mathcal{C}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$; each member of $\mathcal{P}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$ can be described with some LHV.

$\mathcal{Q}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$ The set of quantum probability vectors in $\mathcal{C}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$; each member of $\mathcal{Q}_{m_{\mathcal{A}}; m_{\mathcal{B}}}^{n_{\mathcal{A}}; n_{\mathcal{B}}}$ can be realized by some quantum strategy.

$A_{s_a}^{o_a}$ The POVM element associated with the o_a^{th} outcome of Alice's s_a^{th} measurement

$B_{s_b}^{o_b}$ The POVM element associated with the o_b^{th} outcome of Bob's s_b^{th} measurement

$E(A_{s_a}, B_{s_b})$ Correlation function associated with Alice measuring A_{s_a} and Bob measuring B_{s_b} .

$\mathcal{B}(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}})$ Bounded operator acting on the Hilbert space $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$.

$\mathcal{H}^{[k]}$ The Hilbert space associated with the k^{th} subsystem.

$|\Phi_d^+\rangle$ The d -dimensional maximally entangled state.

$\rho_{d \times d}$ The $d \times d$ -dimensional maximally mixed state, i.e., $\rho_{d \times d} = \frac{\mathbb{1}_d \otimes \mathbb{1}_d}{d^2}$.

- $\rho_{\text{W}_d}(p)$ The d -dimensional Werner state.
- $\rho_{\text{I}_d}(p)$ The d -dimensional isotropic state.
- $\mathcal{S}_{\text{LHV}}^{(\text{k})}$ Functional form of the Bell inequality labeled by “k”.
- $\mathcal{S}_{\text{QM}}^{(\text{k})}(\rho)$ Maximal expectation value of the Bell operator derived from the Bell inequality “k” with respect to the quantum state ρ .
- \mathcal{B}_k The Bell operator derived from the Bell inequality “k”.
- $\langle \mathcal{B}_k \rangle_\rho$ Expectation value of the Bell operator \mathcal{B}_k with respect to a quantum state ρ .
- $\lfloor a \rfloor$ The largest integer smaller than a
- $[M]_{i,j}$ The (i, j) entry of a matrix M .
- M^{T} The transpose of M .
- M^{T_k} The partial transpose of M with respect to the k^{th} subsystem.
- $\text{tr}(M)$ The trace of M .
- $\text{tr}_{\mathcal{A}}(M)$ The partial trace of M over subsystem \mathcal{A} .
- $\|M\|$ The trace norm of M , i.e., the sum of the absolute value of M ’s eigenvalues.
- Π Projector, i.e., $\Pi^2 = \Pi$.
- $\mathbf{1}_d$ The $d \times d$ identity matrix.
- $\mathbf{0}$ The null operator/ zero matrix.
- $\mathbf{0}_n$ The $n \times 1$ null vector.
- $\mathbf{0}_{d_{\mathcal{A}} \times d_{\mathcal{B}}}$ The $d_{\mathcal{A}} \times d_{\mathcal{B}}$ zero matrix.
- $\mathcal{C}_{\text{SLOCC}}^{(\text{CHSH})}$ The set of quantum states not violating the Bell-CHSH inequality even after arbitrary local filtering operations.